AD_____

Award Number: DAMD17-02-2-0011

TITLE: Structural Studies on Intact Clostridium Botulinum Neurotoxins Complexed with Inhibitors Leading to Drug Design

PRINCIPAL INVESTIGATOR: Subramanyam Swaminathan, Ph.D.

CONTRACTING ORGANIZATION: Brookhaven National Laboratories
Upton, NY 11973

REPORT DATE: February 2006

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-02-2006	Annual	28 Jan 2005 – 27 Jan 2006

4. TITLE AND SUBTITLE
Structural Studies on Intact Clostridium Botulinum Neurotoxins Complexed with Inhibitors Leading to Drug Design

5a. CONTRACT NUMBER

5b. GRANT NUMBER
DAMD17-02-2-0011

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Subramanyam Swaminathan, Ph.D.

E-mail: swami@bnl.gov

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Brookhaven National Laboratories
Upton, NY 11973

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

In this fourth annual report we present our progress on several fronts. We have identified a number of inactive mutants of BoNT/E-LC which may be potentially used as vaccine candidates. Crystal structures of several catalytic domains of clostridial neurotoxins have been determined and it has been shown that there is a commonality of the architecture of the active site. Accordingly, it must be possible to design a common drug for all of them. We have determined the structure of the C fragment of botulinum neurotoxin type B and found that the N-terminal helix reorients. Virtual screening with small molecule library has been successfully carried out with BoNT/E-LC. The C fragment of BoNT/A has been successfully cloned and expressed.

15. SUBJECT TERMS
Clostridium, botulinum, neurotoxin, zinc chelators, inhibitors, macromolecular crystallography, 3D structure

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					USAMRMC
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U	UU	54	

Table of Contents

Structural Studies on Intact *Clostridium botulinum* Neurotoxins Complexed with Inhibitors Leading to Drug Design
Annual Report for the Period ending January 2005

Introduction

The overall goal of this project is to identify potential inhibitors for botulinum neurotoxins via macromolecular crystallography. Macromolecular crystallography plays a major role in drug design and uses three-dimensional structures of proteins to identify small molecules that will block the activity by shielding the active or binding sites. Botulinum neurotoxins follow a four-step process, viz., binding to neuronal cells, internalization into vesicles, translocation of catalytic domain into cytosol through the endosomal membrane and the cleavage of one of the three target proteins responsible for neurotransmission. Blocking anyone of these steps will deactivate the protein. In our project we have chosen two of them: blocking the binding site on the binding domain and shielding the active site to prevent catalytic action. A two-pronged approach is being used. We are trying either small molecules which could chelate the active-site zinc or peptide mimics of the substrates to block the active site residues. The general approach is to study the crystal structure of the toxin in complex with a potential inhibitor via x-ray crystallography and then analyze the interactions between the inhibitor and the protein.

Body

(1) Studies on the binding domain of C. neurotoxins

(a) C. botulinum neurotoxin type B:

In the last annual report, we had reported the progress on our work on the crystal structure of the binding domain of botulinum neurotoxin type B. This work was done in collaboration with Drs Len Smith and Syed Ahmed of USAMRIID. The study has been successfully completed and published (*1*). An electronic version of the reprint is attached to this report (Appendix 1). As described in the last report, the structure is very similar to the binding domain in the holotoxin with some changes in the loop regions and the N

terminal helix. The binding site of ganglioside as identified in the holotoxin is unchanged and accordingly this structure can be used to identify small molecules that will block the binding site of botulinum neurotoxin type B. One of the major goals in the next year is to test the molecules that have been identified with tetanus toxin. We will combine both virtual screening (DOCKing) and x-ray crystallography.

(b) C. botulinum neurotoxin type A:

As a continuation of this project on the binding domains of botulinum neurotoxins, we have successfully cloned and expressed the binding domain of botulinum neurotoxin type A. The binding domain corresponding to aa 871-1295 has been cloned into pET28b vector. This will be crystallized and structure determined. The structure will be used to identify small molecules to block the binding site.

(2) Studies with C. neurotoxin catalytic domains

In this section we will describe our work on the catalytic domain of various serotypes of botulinum neurotoxins. This was done to understand the mode of action of all serotypes which have different substrates. The study showed that the active site geometry has a common pattern and it may be possible to develop a common inhibitor to all of them.

(a) Clostridium botulinum Type E catalytic domain

In the last annual report we had reported our progress on expressing BoNT/E catalytic domain and its structure determination (Fig. 1). A paper was published on the crystal structure of the light chain and a PDF file was submitted with the last annual report. We had also described our work on other mutants of BoNT/E catalytic domain. This was carried out for two reasons. (1) To identify residues that are important for the catalytic activity and then design inhibitors to block these residues, and (2) to understand the importance of each residue for the activity and thereby creating a mutant which would invoke the same antibody response but will be unable to cleave the substrate. Such a mutant may eventually be a valuable and potential candidate for vaccine development. We have now completed that work and a paper has been published in Biochemistry journal. An electronic version of the reprint is attached (Appendix 2). In summary, we have determined structures of the following mutants: Arg347Ala, Tyr350Ala,

Glu335Gln, Glu335Ala, and a triple mutant. The study revealed the role of the nucleophilic water and the charge on Glu212/Glu3335 on the catalytic activity (*2*). It also provides a method of developing a novel genetically-modified-protein vaccine. It suggests that a sensitive inhibitor could be developed if Glu212/Glu335 is blocked or covalently modified and may well become a common drug for all serotypes. Also, since they are spatially not far apart, it should be possible to design an inhibitor which will block both by interacting with both of them. Since the conformation of the mutant is almost identical to the wild-type protein and it effectively binds to the substrate, the mutants offer themselves as candidates for studying the enzyme-substrate complex without the substrate being cleaved.

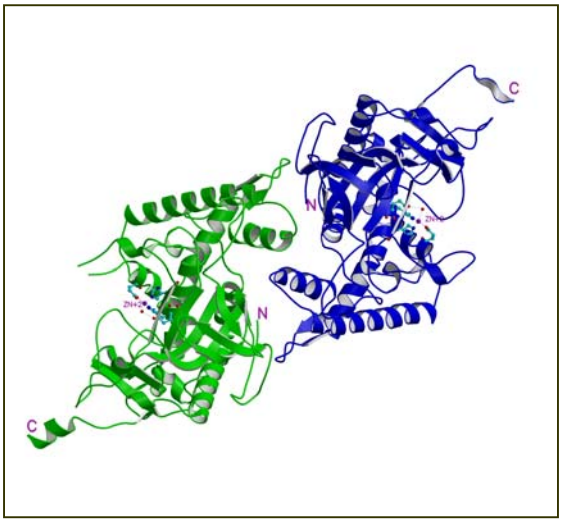

Fig. 1. Ribbons representation of BoNT/E-LC dimer. The two monomers are in green and blue. The active site zinc and the coordinating protein residues are in ball and stick model.

(b) Studies on C. neurotoxin type F catalytic domain:

The BoNT/F catalytic domain was cloned, expressed, purified and the enzymatic activity tested on its substrate VAMP-2. The protein crystallized in two crystal forms and the crystal structures of both forms were determined to a very high resolution by the molecular replacement method. The BoNT/F-LC structure is similar to BoNT/A, /B, /E and TeNT LCs with respect to the core conformation involving the helices and strands

though large variations exist in their loop conformations. The active site geometry is very similar to other botulinum toxins. The active site zinc is coordinated by His227 NE2, His231 NE2, Glu266OE1, and a nucleophilic water molecule (*3*). The latter is kept in its position via Glu228. Counterparts of the remaining depicted active site residues are present in analog arrangement in all clostridial neurotoxins crystallized to date. This argues for the functional similarity of these residues. Results from this study have been published in Biochemistry (Appendix 3). We have identified the S1' site in BoNT/F-LC (Fig. 2).

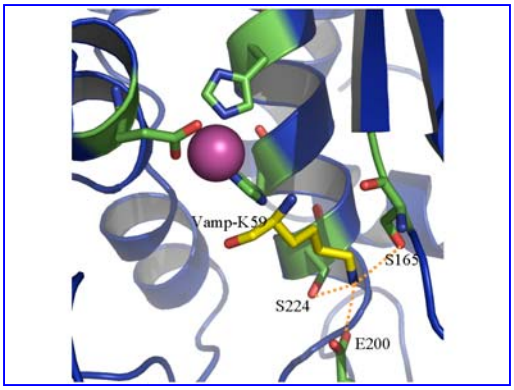

Figure 2. Close up view of the active site and S1' subsite. The active site zinc is shown in magenta and the residues involved in the active and S1' subsite are shown in stick model. The P1' residue Lys59 is shown in yellow. The three potential hydrogen bonds are shown as dashed lined.

(c) Mercury compounds as inhibitors of botulinum neurotoxins:

The role of mercurial compounds on the activity of BoNTs has been investigated with respect to the interchain disulfide bridge connecting the heavy and light chains (*4*). However, the same compound had different effects on various serotypes. Ahmed et al. also have shown that similar mercurial compounds completely abolish the activity of BoNT/A-LC at 10 – 50 μM concentrations (*5*) and this is attributed to the modification of the thiol group of Cys164 which is in the vicinity of the active site. We chose to investigate the effect of mercury compound on BoNT/F since there are two adjacent cysteines (Cys166 and Cys167; Fig. 3) in the beta strand that would form a beta sheet with VAMP according to our enzyme-substrate complex model. Cys167 of BoNT/F and Cys164 of BoNT/A are conserved. In addition, Cys364 is close to Arg365 and Tyr368. The three mercury compounds we tried on BoNT/F-LC had varying effects in inhibiting

the catalytic activity. As has been pointed out (*4*), this could be because of the modification of thiol groups or histidines near the active site or both. Future crystallographic investigations may help in understanding the role of mercury compounds in inhibiting the catalytic activity.

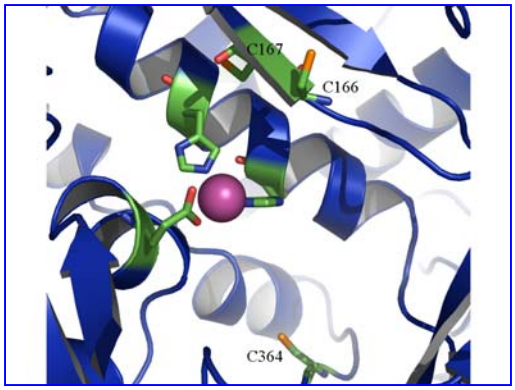

Figure 3. The three cysteines (166, 167 and 364) within 11 Å from the active site zinc are shown. Cys364 is close to Arg365 and Tyr368 which stabilize the transition state. The fourth cysteine (Cys429) was not modeled in the structure.

(d) Structural studies on C. neurotoxin type B catalytic domain:

The available crystal structure of BoNT/B-LC is that of a truncated protein and lacks 20 C-terminal residues (*6*). Since it is believed that the C-terminal region makes contacts with the substrate, we crystallized full-length type B catalytic domain and determined the structure by the molecular replacement method.

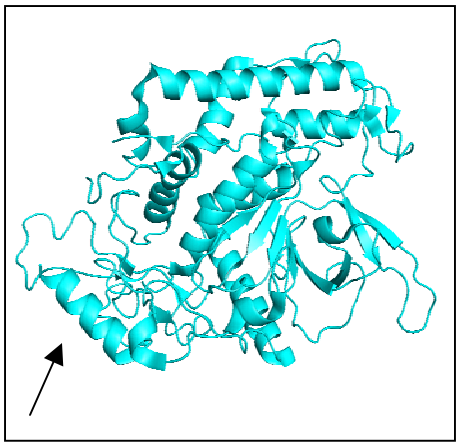

Figure 4. The catalytic domain of BoNT/B. The C-terminal helix is shown by an arrow mark.

Unlike other botulinum catalytic domain structures, the whole molecule was ordered (except for two short loops) and the electron density was interpretable from N to C terminals. The C-terminal region took a helical conformation (shown by an arrow in Figure 4) in contrast to the holotoxin structure (*7*). This change in conformation on separation from the rest of the molecule may have some role in its localization in the cytosol. Since this is a full-length LC and high-resolution structure we will be using this crystal form for inhibitor studies.

(e) Structural studies on Tetanus neurotoxin catalytic domain:

Though we are mainly interested in botulinum neurotoxins, as a part of our studies on clostridial neurotoxins, we determined the crystal structure of TeNT-LC also (*8*). As expected, the fold was similar to botulinum neurotoxins (Fig. 5).

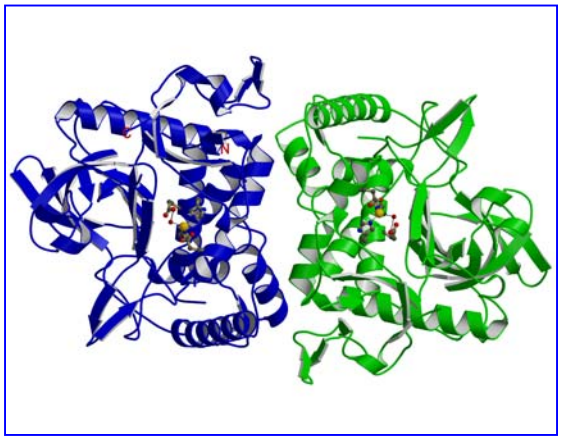

Figure 5. Ribbons representation of TeNT-LC. The two monomers forming a dimer in the crystal structure are shown in blue and green. Active site residues are shown in ball and stick model.

This study was taken up mainly to understand the difference between BoNT/B and TeNT. These two have the same substrate but cleave it at different peptide bonds. Moreover, the optimum length of the substrate required for its cleavage is different. Our aim was to investigate the reason for this and is discussed in the paper published in Biochemistry. An electronic version of the reprint is attached (Appendix 4).

(f) Virtual Screening of small molecule database:

Initially, we considered ZINC database (Zinc Is Not Commercial) for virtual screening (*9*). Though it contains more than 72,000 compounds, we restricted to 50,000 compounds for our preliminary studies. At first we checked virtual screening procedure

with BoNT/E-LC. BoNT/E-LC was prepared for DOCK program (*10*) with suitable charges, etc. This was done with local scripts written in our laboratory and we used SGI origin 300 for computations. Since the code that we used was not parallelized, it took more than 30 days to analyze 50,000 compounds with BoNT/E-LC. We chose this serotype since we had a high-resolution x-ray structure. This procedure filtered out about 31,000 compounds and only 19,000 could be docked with reasonable energies. Thus the virtual screening procedure had filtered out 62% of the compounds in the database. Further work is on progress.

Key Research Accomplishments

- Crystal structures of BoNT/B, BoNT/F and TeNT catalytic domains have been determined helping us to understand the commonality in the mechanism of action of clostridial neurotoxins.
- The binding site of mercury compounds to inhibit the catalytic action has been identified. This may open up a new method of inhibiting the toxicity of clostridial neurotoxins.
- A subset of small molecules that may inhibit the catalytic activity of BoNT/E has been identified.
- BoNT/A binding domain has been cloned and expressed.
- The crystal structure of BoNT/B C fragment has been determined.

Reportable outcomes

1. Rakhi Agarwal, T. Binz and S. Swaminathan. Analysis of active site residues of botulinum neurotoxin E by mutational, functional and structural studies: Glu335Gln is an apoenzyme. *Biochemistry*, 2005, **44**, 8291-8302.

2. Rakhi Agarwal, T. Binz and S. Swaminathan. Structural analysis of botulinum neurotoxin serotype F light chain: Implications on substrate and inhibitor binding. *Biochemistry*, 2005, **44**, 11758-11765.

3. Krishnamurthy N. Rao, D. Kumaran, T. Binz and S. Swaminathan. Structural analysis of the catalytic domain of tetanus neurotoxin. *Toxicon*, 2005, **45**, 929-939.

4. S. Jayaraman, S. Eswaramoorthy, S. A. Ahmed, L. A. Smith and S. Swaminathan. N-terminal helix reorients in recombinant C-fragment of *Clostridium botulinum* type B. *Biochem. Biophys. Res. Commun.*, 2005, **30**, 97-103.

Conclusions

In our studies we have shown that there is common architecture at the active site of clostridial neurotoxins which may be exploited to design a common drug for all serotypes. We have also shown residues important for the catalytic activity of BoNT/E and the possibility of blocking these residues to inhibit them. Some of the mutants of BoNT/E which are completely devoid of activity, may be used as genetically modified vaccines.

Plans for the next quarter:

We have identified a few compounds which may be good inhibitors of BoNT/E. Crystal structures of theses compounds in complex with the catalytic domain will be undertaken and completed. The work on virtual screening will be completed. Crystal structure of the binding domain of BoNT/A C fragment will be completed and used for virtual screening of small molecules to block the binding site. Also structural work on BoNT/B-inhibitor complex will be continued.

Personnel in the Project

1. S. Swaminathan (PI)	Scientist	30% effort
2. S. Eswaramoorthy	Associate Scientist	25% effort
3. R. Agarwal	Research Associate	75% effort

Reference:

1. Jayaraman, S., Eswaramoorthy, S., Ahmed, S. A., Smith, L. A., and Swaminathan, S. (2005) N-terminal helix reorients in recombinant C-fragment of Clostridium botulinum type B, *Biochem Biophys Res Commun 330*, 97-103.
2. Agarwal, R., Binz, T., and Swaminathan, S. (2005) Analysis of active site residues of botulinum neurotoxin E by mutational, functional and structural studies: Glu335Gln is an apoenzyme, *Biochemistry 44*, 8291-8302.
3. Agarwal, R., Binz, T., and Swaminathan, S. (2005) Structural analysis of botulinum neurotoxin serotype f light chain: implications on substrate binding and inhibitor design, *Biochemistry 44*, 11758-65.

4. Simpson, L. L., Maksymowych, A. B., Park, J.-B., and Bora, R. S. (2004) The role of the interchain disulfide bond in governing the pharmacological actions of botulinum toxin, *J. Pharmacol. Exp. Ther. 308*, 857-864.
5. Ahmed, S. A., and Smith, L. A. (2000) Light chain of botulinum A neurotoxin expressed as an inclusion body from a synthetic gene is catalytically and functionally active, *J. Protein Chem. 19*, 475-487.
6. Hanson, M. A., and Stevens, R. C. (2000) Cocrystal structure of synaptobrevin-ll bound to botulinum neurotoxin type B at 2.0 A resolution, *Nature Struct. Biol. 7*, 687-692.
7. Swaminathan, S., and Eswaramoorthy, S. (2000) Structural analysis of the catalytic and binding sites of *Clostridium botulinum* neurotoxin B, *Nature Struct. Biol. 7*, 693-699.
8. Rao, K. N., Kumaran, D., Binz, T., and Swaminathan, S. (2005) Structural studies on the catalytic domain of clostridial tetanus toxin, *Toxicon 45*, 929-939.
9. Irwin, J. J., and Shoichet, B. K. (2005) ZINC - A free database of commercially available compounds for virtual screening, *J. Chem. Inf. Model 45*, 177-182.
10. Kuntz, I. D. (1992) Structure-based strategies for drug design and discovery, *Science 257*, 1078-1082.

Appendix 1

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 330 (2005) 97–103
BBRC
www.elsevier.com/locate/ybbrc

N-terminal helix reorients in recombinant C-fragment of *Clostridium botulinum* type B

Seetharaman Jayaraman [a], Subramaniam Eswaramoorthy [a], S. Ashraf Ahmed [b], Leonard A. Smith [b], Subramanyam Swaminathan [a,*]

[a] *Biology Department, Brookhaven National Laboratory, Upton, NY 11973, USA*
[b] *Department of Immunology and Molecular Biology, Division of Toxicology, United States Army Medical Research Institute of Infectious Diseases, Fort Detrick, MD 21702, USA*

Received 3 February 2005

Abstract

Botulinum neurotoxins comprise seven distinct serotypes (A–G) produced by *Clostridium botulinum*. The crystal structure of the binding domain of the botulinum neurotoxin type B (BBHc) has been determined to 2 Å resolution. The overall structure of BBHc is well ordered and similar to that of the binding domain of the holotoxin. However, significant structural changes occur at what would be the interface of translocation and binding domains of the holotoxin. The loop 911–924 shows a maximum displacement of 14.8 Å at the farthest point. The N-terminal helix reorients and moves by 19.5 Å from its original position. BBHc is compared with the binding domain of the holotoxin of botulinum type A and B, and the tetanus C-fragment to characterize the heavy chain-carbohydrate interactions. The probable reasons for different binding affinity of botulinum and tetanus toxins are discussed.
Published by Elsevier Inc.

Keywords: *Clostridium botulinum*; Binding domain; X-ray crystallography; Three-dimensional structure; Gangliosides

Botulinum neurotoxin (BoNT) is a member of the clostridial neurotoxin family (CNT), which includes seven botulinum neurotoxins (A to G) and tetanus (TeNT) toxin. CNTs are the most potent toxins having selective high affinity binding to neurons. BoNTs act at the neuromuscular junction causing flaccid paralysis while TeNT acts on peripheral nervous system causing spastic paralysis. These toxins are synthesized as inactive single chain proteins of ~150 kDa and released as di-chains cleaved either by endogenous or exogenous proteases. The di-chain is composed of a heavy chain (100 kDa) and a light chain (50 kDa) associated via both inter-chain interactions and the conserved disulfide bond between heavy and light chains [1,2].

Polysialogangliosides bind CNTs on the C-terminal domain of heavy chain [3]. Many experimental studies have predicted the location, binding specificity, and affinity for ganglioside binding on the C-terminal subdomain of CNTs [4–10]. A double receptor model has been suggested for CNTs, binding to neuronal cells and might involve both the subdomains [11]. Low (K_d in nM range) and high (K_d in sub-nM) affinity receptors for CNTs have been identified and experimental evidences prove that for high affinity binding, both lipid and protein receptors are essential [3]. It has been shown that the C-terminal domain of BoNT and TeNT (TeHc) is necessary and sufficient for binding and internalization of these proteins into neurons [12].

The crystal structures of BoNT/A and B show that the molecules have three distinct structural domains corresponding to three functional domains; (i) the catalytic

* Corresponding author. Fax: +1 631 344 3407.
E-mail address: swami@bnl.gov (S. Swaminathan).

0006-291X/$ - see front matter. Published by Elsevier Inc.
doi:10.1016/j.bbrc.2005.02.123

domain containing Zn-metalloprotease active site; (ii) a translocation domain containing a long loop (residues 481–532), called the belt region that wraps around the catalytic domain; and (iii) a binding domain. The amino acid sequence of the binding domains is well conserved among all clostridial neurotoxins, suggesting that they may have a closely similar structure [13,14]. The crystal structures of the binding domain of TeHc, BoNT/A and B show two distinct subdomains within the binding domain, N- and C-terminal domains [15–17]. The C-terminal β-trefoil domain contains folds related to those present in several other proteins involved in recognition and binding functions, such as various trypsin inhibitors, and the N-terminal domain contains a lentil lectin-like jelly-roll motif [18].

A number of studies have illustrated the importance of the C-terminal region of the β-trefoil domain, however the lectin-like carbohydrate binding N-terminal domain has received less attention. The crystal structures of BoNT/B with doxorubicin and sialyllactose [17] show that both bind at the same site and crystal structures of TeHc with sialic acid, *N*-acetylgalactosamine, GT1b, galactose, disialyllactose, and a tripeptide (YEW) show that they bind in four different sites. Interestingly, all these binding sites are in the C-terminal domain [8,10,19,20]. Though N-terminal domain is similar to lectin carbohydrate binding domain, so far there has been no evidence of sugars binding in this region.

BoNT and TeNT act by blocking neurotransmitter release. Their mechanism of action on target nerve cells consists of binding, internalization, and membrane translocation by heavy chain and enzymatic target cleavage by light chain [1]. Separation of the domains from the holotoxin is expected to result in significant structural changes since the hydrophobic contacts at the interface would be lost [21]. Binding domains of botulinum toxins are being actively pursued as potential vaccine candidates. Since a potential vaccine candidate should evoke the same kind of immune response as the binding domain in the holotoxin, no large structural change should occur. The structural work was undertaken to test whether the structural integrity is maintained when the binding domain is separated from the holotoxin. Here, we report the crystal structure of the binding domain of BoNT/B (BBHc) and compare, with the holotoxin binding domain of BoNT/A and B, and TeHc.

Experimental

BBHc containing residues 853–1290 of *Clostridium botulinum* type B, Danish strain was expressed from a synthetic gene in *Pichia pastoris* and purified as described [22]. The protein in 0.05 M $NaPO_4$ buffer at pH 7.5, with, initial concentration of 0.130 mg/ml was dialyzed against 0.02 M Hepes and 0.1 M NaCl at pH 7.5, for 48 h in four changes at 4 °C. The protein was then concentrated to 6.5 mg/ml. Crystallization screening was carried out with Hampton Crystal Screen, and two or three conditions were selected. The best condition was with ammonium acetate which gave very small crystals and was later successfully replaced by ammonium formate. Sitting drop crystallization was set up with a droplet containing a mixture of 1 μl protein, 1 μl of 60–100 mM ammonium formate, and 1 μl of 30% PEG 6000. Different sugars and additives were tried to improve the quality of the crystal. Addition of 0.3 μl of the additive 2 M non-detergent Sulfo-betaine 201 gave better crystals. The reservoir solution consists of 20% PEG6000 and 0.1 M Hepes at pH 7. Rod-shaped crystals of dimension $0.05 \times 0.06 \times 0.08$ mm appeared after a week and diffracted to 1.9 Å. Crystals are in space group $P2_1$ with cell dimensions $a = 68.66$, $b = 78.81$, $c = 88.72$ Å and $\beta = 103.03°$. Matthews coefficient (V_m) was calculated to be 2.37 $Å^3$/Da based on two molecules per asymmetric unit [23]. The two molecules in the asymmetric unit are related by a twofold at 1/4 along *c*-axis, parallel and halfway between crystallographic twofold screws.

Data were collected at liquid nitrogen temperature at X25 beam line of the NSLS, Brookhaven National Laboratory with Q315 detector. An oscillation range of 1° was used for each data frame with the crystal to detector distance of 280 mm and $\lambda = 1.01$ Å. Data were processed by HKL2000 [24].

The structure was determined by the molecular replacement method using the poly-alanine model of the binding domain of the previously determined BoNT/B structure (PDB id: 1ERW) as the search model [25]. Loops were truncated in the model to avoid model bias. After initial rigid body refinement, side chains were included in the model. The model was refined by slow cool simulated annealing method with CNS and in each step, $2F_{obs} - F_{calc}$ and $F_{obs} - F_{calc}$ maps were calculated to check and improve the model in the density maps using the program O [26,27]. Cycles of rebuilding and refinement were

Table 1
Data reduction and refinement statistics

Space group	$P2_1$
Number of molecules per asymmetric unit	2
Unit cell	$a = 68.66$, $b = 78.81$, $c = 88.72$ Å and $\beta = 103.03°$
Resolution (Å)	2.0
$I/\sigma(I)$	11.0
Redundancy	4.0 (3.2)[a]
No. of unique reflections	59,503 (4723)
Completeness	96 (77)
R_{merge}[b]	0.09 (0.48)
R-factor[c]	0.23
R_{free}[d]	0.28
No. of protein atoms	3564
No. of water molecules	431
Average *B*-factor (Å)	
Protein	32.3
Water molecules	38.1
RMSD	
Bond lengths (Å)	0.01
Bond angles (°)	1.50

[a] Values in parentheses are for the highest resolution shell, 2.07–2.0 Å.

[b] $R_{merge} = \sum|I_i - \langle I \rangle|/\sum|I_i|$, where I_i is the intensity of the *i*th measurement, and $\langle I \rangle$ is the mean intensity for that reflection.

[c] *R*-factor $= \sum|F_{obs} - F_{calc}|/\sum|F_{obs}|$, where F_{calc} and F_{obs} are the calculated and observed structure factor amplitudes, respectively.

[d] R_{free} = as for R_{cryst}, but for 10.0% of the total reflections chosen at random and omitted from refinement.

continued until the convergence of *R*-factor and *R*-free. A total of 430 water molecules were located from the difference Fourier maps and were included in subsequent rounds of refinement. The model is complete except that 1150–1156 loop region is disordered. Data reduction and refinement statistics are listed in Table 1 (PDB id: 1Z0H).

Results and discussion

BBHc structure

The structure of the BBHc has been determined to 2.0 Å resolution. The molecule has maximum dimension of approximately $33 \times 42 \times 80$ Å^3 and has 438 residues from Asn853 to Glu1290. The molecule comprises two similar sized subdomains (C- and N-terminal domains) which have limited surface interactions with each other. The tertiary structure suggests that the two domains fold independently. The N-terminal domain has 16 β-strands (which includes a set of seven-stranded antiparallel β-sheets) and three α-helices, and the C-terminal domain has 16 β-strands and one α-helix (Fig. 1). The overall structure of BBHC is well ordered and globally similar to that of the binding domain of holotoxin structure, except for the orientation of two loops located on the surface of the molecule and the N-terminal helix [17]. Analysis of the model using PROCHECK shows that all non-proline and non-glycine residues are found in the most favored or additionally allowed region of the Ramachandran plot [28].

Fig. 1. Ribbon diagram showing the tertiary structure of the binding domain of BBHc.

Table 2
Binding domain, translocation domain interactions in holotoxin BoNT/B

Binding domain	Translocation domain	Distance (Å)
Tyr 855 OH	Lys 784 NZ	2.94
Asn 866 ND2	Asp 615 OD2	2.96
Arg 868 NH1	Ile 613 O	3.31
Arg 868 NH2	Asp 615 OD2	3.30
	Ile 613 O	3.43
Lys 870 NZ	Asn 640 OD1	3.28
	Lys 785 NZ	2.68
Tyr 880 OH	Lys 784 NZ	3.19
Asn 915 OD1	Glu 792 N	3.51

Separation of BD

The binding domain and the rest of the holotoxin are linked through the N-terminal α-helix of the binding domain. In the holotoxin structure, the binding domain is tilted away from the central translocation domain and makes minimal contact with it. The two domains interact via nine hydrogen bonds (Table 2). There are 13 other possible hydrophilic contacts within 4.8 Å as calculated by the program CONTACT from the CCP4 suite of programs (Collaborative Computational Project, No. 4, 1994). Water molecules in the interface play an important role in binding these domains [17]. In addition, numerous hydrophobic contacts are observed between the two domains involving the N-terminal helix and the loop at the interface (aa 911–924). The separation of binding domain causes the loss of these interactions resulting in major loop movements.

BBHc and BoNT/B

Comparison of the isolated binding domain and the binding domain of holotoxin was carried out to evaluate structural changes. The superposition of BBHc on BoNT/B holotoxin binding domain (Fig. 2) shows that the N-terminal helix and loop regions on the surface of the molecule do not superimpose, particularly the loop at translocation and binding domain interface. Overall, one loop in the N-terminal domain and two loops in the C-terminal domain show significant displacements. (i) In the holotoxin, the N-terminal helix (aa 846–856) of binding domain is close and almost orthogonal to long helices in the translocation domain and makes extensive contacts with the translocation domain helices (Table 2). The N-terminal helix bridges the two domains. After the domain separation, it loses all its interactions and moves 19.5 Å away from its position changing its orientation. Its orientation is similar to

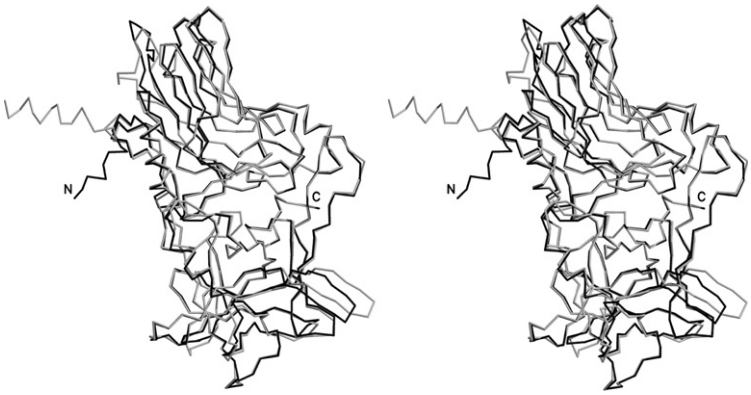

Fig. 2. Stereo view of superposition of isolated binding domain with the binding domain of the holotoxin. Black-isolated binding domain and gray-holotoxin.

those of the corresponding helices in the binding domain of BoNT/A and TeHc. However, BoNT/A is a holotoxin whereas TeHc is binding domain only. (ii) The loop (aa 911–924) at the interface of translocation and binding domain which makes direct hydrogen bonds with the translocation domain is displaced by 14.7 Å in the absence of these interactions. The buried hydrophobic residues at this region are exposed in BBHc. The solvent accessibility of the loop in the absence of translocation domain increases. (iii) One of the loops in the C-terminal region (aa 1245–1250) moves away from its position by 7.5 Å. This region binds sialyllactose and doxorubicin in the cleft between His1240 and Trp1261 [8,17]. Though the extended loop shows a displacement, the binding pocket which is at the beginning of the loop region maintains its shape and size. This loop movement may not affect any binding in this region. The binding site comparison shows a general widening of the ligand binding cleft at the extended binding site region that can facilitate the binding.

As the binding domain shows significant changes, its counterpart at translocation domain is expected to show similar changes particularly at the interface region on separating it from holotoxin. The separation of catalytic domain of BoNT/B showed similar rearrangement of the loops in the active site exposing the active site for binding [29]. Such changes are also observed in BoNT/E and TeNT (KN Rao, unpublished) catalytic domains [30].

Comparison with BoNT/A and TeNT binding domain

The binding domains of BoNT/B, BoNT/A, and TeNT have 438, 430, and 441 residues, respectively. Fig. 3A shows the superposition of the Cα atoms of BBHc (green), binding domain of BoNT/A (blue) and TeHc (red). The rms deviation between BBHc and BoNT/A is 1.4 Å for 349 Cα atoms, and between BBHc and TeHc is 1.3 Å for 346 Cα atoms. The three models provide general information on the heavy chain-carbohydrate interactions. The superposition shows that the core β-strand regions superimpose fairly well whereas loop regions do not.

Hc–carbohydrate interaction

There are two BoNT/B complex structures and six TeHc complex structures available for comparison, to understand the Hc–carbohydrate interaction of the binding domain. BoNT/B and TeNT have a common binding site, Site-1 (Fig. 3A). This site in BoNT/B binds sialyllactose and doxorubicin (His1240 and Tyr1262 region), and in TeNT binds lactose and part of GT1b (Gal-GalNac) (His1271 and Tyr1290 region) [10,19]. The binding pocket has the same size and shape in both molecules as this region is structurally and sequentially well conserved. Various biochemical studies indicated that this region of β-trefoil subdomain of clostridial neurotoxins was involved in ganglioside binding. Site-directed mutagenesis showed that Tyr1290 plays a key role in binding of ganglioside in TeNT [6]. This tyrosine appears to be conserved throughout CNTs: Tyr1290 in TeNT, Tyr1266 in BoNT/A, and Tyr1262 in BoNT/B. Ganglioside-photoaffinity ligand study implicated that His1293 of this region in TeNT is important for ganglioside binding [31]. Gln1269 in BoNT/A and Glu1265 in BoNT/B can play a similar role as His1293 in TeNT with a similar stacking of the side chain of Trp1289, which is used to maintain the binding pocket in proper shape for the binding to occur. Fluorescence experiments suggested that Trp1265 must be part of ganglioside binding site in BoNT/A [32]. Trp1261 in BoNT/B was shown to be involved in binding sialyllactose in this region [17]. As BoNT/A has the same sequence and structural motif, it is also expected to bind ganglioside in this region. This site is a conserved site and represents the general binding pocket of clostridial neurotoxins. As there is no antidote available for botulinum at present,

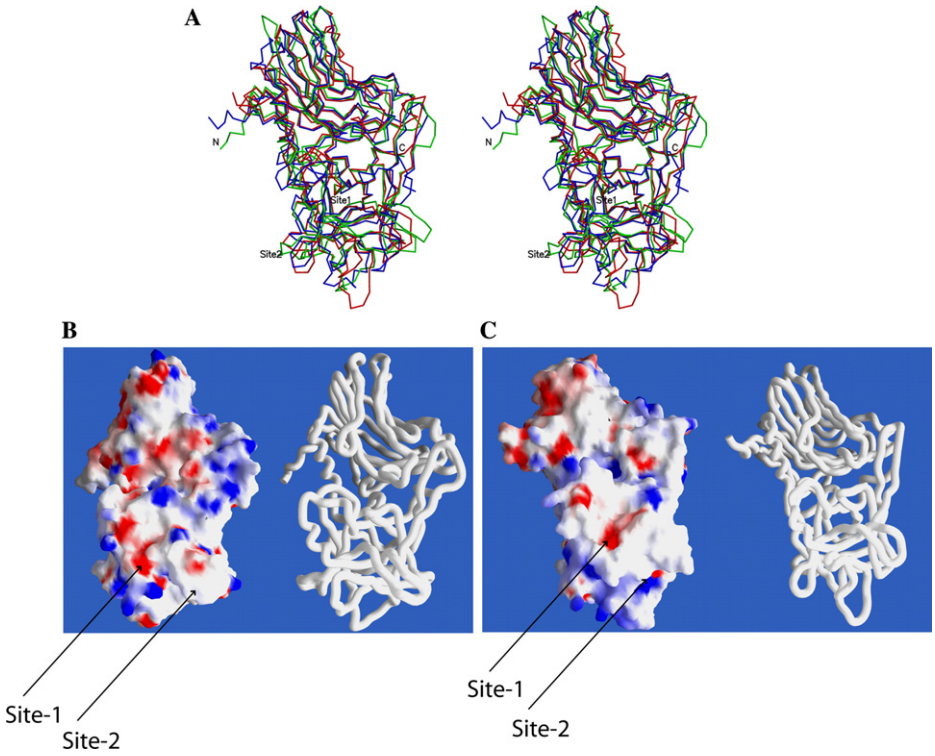

Fig. 3. (A) Stereo view of superposition of the Cα atoms of BBHc (green), TeHc (red), and binding domains of BoNT/A (blue). (B) Electrostatic potential of the recombinant binding domain of BoNT/B. Blue and red represent the positive and negative electrostatic potentials, respectively. Adjacent figure shows the worm representation of the molecule. Site-1 and Site-2 are marked. (C) A similar representation for TeHc.

this site is a potential target for development of therapeutics.

The other important site, Site-2 (Fig. 3A), the region around 1114–1147 of TeHc, binds sialic acid, disialyllactose, YEW, and Gal-NAc part of GT1b [10,19,20]. In this region, there is no structural evidence of any binding in BoNT/A and B. A comparison with BoNT/A and B shows that the length and orientation of the loops forming the binding site are different. The key residues Asp1214, Asn1216, Asp1147, and Tyr1229 that interact with bound ligands in TeHc are on the loop region and are not conserved in BoNT/A and B. This is the unique and most common sugar-binding site for TeNT but differs structurally and sequentially from BoNT/A and B.

The electrostatic potential surface shows the similarity and dissimilarity among the binding sites in the binding domain of BoNT/B and TeHc (Fig. 3B) [33]. The lactose and Gal-GalNAc binding site (Site-2) of TeHc is formed by the side chains of Asp1214, Asn1216, Arg1226, and Tyr1229; this site in TeHC is formed by charged residues while in BoNT/B they are mostly hydrophobic (This region in BoNT/B is made of side chains of Phe1193, Val1117, Leu1194, Val1227, Pro1116, Phe1242, and Phe1203). The electrostatic potential surface at this site is positive in TeHc and neutral in BoNT/B. The shape and electrostatic potential properties of lactose binding site (Site-1) of TeHc (formed by residues Asn1220, Asp1222, Thr1270, Ser1287, Trp1289, Tyr1290, and Gly1300) and sialyllactose-binding site of BoNT/B (formed by residues His1240, Try1262, Glu1189, and Glu1188) appear to be similar. The bottom of the pocket shows a highly hydrophilic patch in TeHc and BoNT/B formed by the structurally equivalent residues. In TeHc, this includes the side chain oxygen of residues Asp1222, Ser1287, and Tyr1290, and the main chain carbonyl oxygen of Thr1270. These residues play a key role in binding the galactose unit of ganglioside [10]. The corresponding residues, Ser1287, Tyr1262 in BoNT/B, play a key role in binding sialyllactose in this region [17]. Tyr1262 of BoNT/B makes two strong hydrogen bonds with the silalic acid part of sialyllactose agreeing with the results of Sutton et al. [6] indicating that Tyr1262 plays an important role in ganglioside binding.

Conclusion

The structure of BBHc shows that the separation of the binding domain from the rest of holotoxin results in large displacements of the loop and the N-terminal helix connecting the translocation and binding domains. This is probably the reason why the C-fragment is a very poor competitor of the toxin for receptor binding. There are no significant changes in the sialyllactose-

binding site between BBHc and BoNT/B-binding domain. Comparison of electrostatic charge distribution of BBHc and TeHc at the binding sites explains the probable reason for the difference in binding affinity between the two.

Acknowledgments

The research was supported by the U.S Army Medical Research Acquisition Activity (Award No. DAMD17-02-2-0011) under DOE Prime Contract No. DE-AC02-98CH10886 with Brookhaven National Laboratory. We thank Dr. D. Kumaran for helpful discussions and Dr. M. Becker for providing beam time at the National Synchrotron Light Source, Brookhaven National Laboratory.

References

[1] G. Schiavo, M. Matteoli, C. Montecucco, Neurotoxins affecting neuroexocytosis, Physiol. Rev. 80 (2000) 717–766.
[2] A. de-Paiva, B. Poulain, G.W. Lawrence, C.C. Shone, L. Tauc, J.O. Dolly, A role for the interchain disulfide or its participating thiols in the internalization of botulinum neurotoxin A revealed by a toxin derivative that binds to ecto-acceptors and inhibits transmitter release intracellularly, J. Biol. Chem. 268 (1993) 20838–20844.
[3] J.L. Halpern, E.A. Neale, Neurospecific binding, internalization, and retrograde axonal transport, Curr. Top. Microbiol. Immunol. 195 (1995) 221–241.
[4] H.A. Louch, E.S. Buczko, M.A. Woody, R.M. Venable, W.F. Vann, Identification of a binding site for ganglioside on the receptor binding domain of tetanus toxin, Biochemistry 41 (2002) 13644–13652.
[5] A. Rummel, S. Bade, J. Alves, H. Bigalke, T. Binz, Two carbohydrate binding sites in the H(CC)-domain of tetanus neurotoxin are required for toxicity, J. Mol. Biol. 326 (2003) 835–847.
[6] J.M. Sutton, O. Chow-Worn, L. Spaven, N.J. Silman, B. Hallis, C.C. Shone, Tyrosine-1290 of tetanus neurotoxin plays a key role in its binding to gangliosides and functional binding to neurones, FEBS Lett. 493 (2001) 45–49.
[7] J.L. Halpern, A. Loftus, Characterization of the receptor-binding domain of tetanus toxin, Nature 268 (1993) 11188–11192.
[8] S. Eswaramoorthy, D. Kumaran, S. Swaminathan, Crystallographic evidence for doxorubicin binding to the receptor-binding site in *Clostridium botulinum* neurotoxin B, Acta Cryst. D57 (2001) 1743–1746.
[9] C. Fotinou, P. Emsley, I. Black, H. Ando, H. Ishida, M. Kiso, K.A. Sinha, N.F. Fairweather, N.W. Isaacs, The crystal structure of tetanus toxin Hc fragment complexed with a synthetic Gt1b analogue suggests cross-linking between ganglioside receptors and the toxin, J. Biol. Chem. 276 (2001) 32274–32281.
[10] P. Emsley, C. Fotinou, I. Black, N.F. Fairweather, I.G. Charles, C. Watts, E. Hewitt, N.W. Isaacs, The structures of the H(C) fragment of tetanus toxin with carbohydrate subunit complexes provide insight into ganglioside binding, J. Biol. Chem. 275 (2000) 8889–8894.
[11] C. Montecucco, How do tetanus and botulinum toxins bind to neuronal membranes? Trends Biochem. Sci. 11 (1986) 314–317.
[12] G. Lalli, J. Herreros, S.L. Osbrone, C. Montecucco, O. Rossetto, G. Schiavo, Functional characterisation of tetanus and botulinum neurotoxins binding domains, J. Cell Sci. 112 (1999) 2715–2724.
[13] N.P. Minton, Molecular genetics of clostridial neurotoxins, Curr. Top. Microbiol. Immunol. 195 (1995) 161–194.
[14] F.J. Lebeda, M.A. Olson, Structural predictions of the channel-forming region of botulinum neurotoxin heavy chain, Toxicon 33 (1995) 559–567.
[15] T.C. Umland, L.M. Wingert, S. Swaminathan, W.F. Furey, J.J. Schmidt, M. Sax, Structure of the receptor binding fragment H_c of tetanus neurotoxin, Nat. Struct. Biol. 4 (1997) 788–792.
[16] D.B. Lacy, W. Tepp, A.C. Cohen, B.R. DasGupta, R.C. Stevens, Crystal structure of botulinum neurotoxin type A and implications for toxicity, Nat. Struct. Biol. 5 (1998) 898–902.
[17] S. Swaminathan, S. Eswaramoorthy, Structural analysis of the catalytic and binding sites of *Clostridium botulinum* neurotoxin B, Nat. Struct. Biol. 7 (2000) 693–699.
[18] J.S. Richardson, The anatomy and taxonomy of protein structure, Adv. Protein Chem. 34 (1981) 167–339.
[19] C. Fotinou, P. Emsley, I. Black, H. Ando, H. Ishida, M. Kiso, K.A. Sinha, N.F. Fairweather, N.W. Isaacs, The crystal structure of tetanus toxin Hc fragment complexed with a synthetic GT1b analogue suggests cross-linking between ganglioside receptors and the toxin, J. Biol. Chem. 276 (2001) 32274–32281.
[20] J. Seetharaman, S. Eswaramoorthy, D. Kumaran, S. Swaminathan, A common binding site for disialyllactose and a tripeptide in the C-fragment of tetanus neurotoxin, Proteins (2005) in press.
[21] D.B. Lacy, R.C. Stevens, Sequence homology and structural analysis of clostridial neurotoxins, J. Mol. Biol. 291 (1999) 1091–1104.
[22] K.J. Potter, M.A. Bevins, E.V. Vassilieva, V.R. Chiruvolu, T. Smith, L.A. Smith, M.M. Meagher, Production and purification of the heavy-chain fragment C of botulinum neurotoxin, serotype B, expressed in the methylotrophic yeast *Pichia pastoris*, Protein Expr. Purif. 13 (1998) 357–365.
[23] B.W. Matthews, Solvent content of protein crystals, J. Mol. Biol. 33 (1968) 491–497.
[24] Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode, Methods Enzymol. 276 (1997) 307–326.
[25] J. Navaza, P. Saludjian, AMoRe: an automated molecular replacement program package, Methods Enzymol. 276 (1997) 581–594.
[26] A.T. Brunger, P.D. Adams, G.M. Clore, W.L. Delano, P. Gros, R.W. Grosse-Kunstleve, J.S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Somonsom, G.L. Warren, Crystallography and NMR system: a new software suite for macromolecular structure determination, Acta Crystallogr. D54 (1998) 905–921.
[27] T.A. Jones, J. Zou, S. Cowtan, M. Kjeldgaard, Improved methods in building protein models in electron density map and the location of errors in these models, Acta Crystallogr. A47 (1991) 110–119.
[28] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality for assessing the accuracy of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[29] M.A. Hanson, R.C. Stevens, Cocrystal structure of synaptobrevin-ll bound to botulinum neurotoxin type B at 2.0 Å resolution, Nat. Struct. Biol. 7 (2000) 687–692.
[30] R. Agarwal, S. Eswaramoorthy, D. Kumaran, T. Binz, S. Swaminathan, Structural analysis of botulinum neurotoxin type E catalytic domain and its mutant Glu212->Gln reveals the pivotal role of the Glu212 carboxylate in the catalytic pathway, Biochemistry 43 (2004) 6637–6644.

[31] R.S. Shapiro, C.D. Specht, B.E. Collins, A.S. Woods, R.J. Cotter, R.L. Schnaar, Identification of a ganglioside recognition domain of tetanus toxin using a novel ganglioside photoaffinity ligand, J. Biol. Chem. 272 (1997) 30380–30386.
[32] Y. Kamata, M. Yoshimoto, S. Kozaki, Interaction between botulinum neurotoxin type A and ganglioside: ganglioside inactivates the neurotoxin and quenches its tryptophan fluorescence, Toxicon 35 (1997) 1337–1340.
[33] A. Nicholls, K.A. Sharp, B. Honig, Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons, Proteins: Struct. Funct. Genet. 11 (1991) 281–296.

Appendix 2

Analysis of Active Site Residues of Botulinum Neurotoxin E by Mutational, Functional, and Structural Studies: Glu335Gln Is an Apoenzyme†

Rakhi Agarwal,‡ Thomas Binz,§ and Subramanyam Swaminathan*,‡

Biology Department, Brookhaven National Laboratory, Upton, New York 11973, and Department of Biochemistry, Medizinische Hochschule Hannover, Hannover, Germany

Received February 10, 2005; Revised Manuscript Received April 18, 2005

ABSTRACT: Clostridial neurotoxins comprising the seven serotypes of botulinum neurotoxins and tetanus neurotoxin are the most potent toxins known to humans. Their potency coupled with their specificity and selectivity underscores the importance in understanding their mechanism of action in order to develop a strategy for designing counter measures against them. To develop an effective vaccine against the toxin, it is imperative to achieve an inactive form of the protein which preserves the overall conformation and immunogenicity. Inactive mutants can be achieved either by targeting active site residues or by modifying the surface charges farther away from the active site. The latter affects the long-range forces such as electrostatic potentials in a subtle way without disturbing the structural integrity of the toxin causing some drastic changes in the activity/environment. Here we report structural and biochemical analysis on several mutations on *Clostridium botulinum* neurotoxin type E light chain with at least two producing dramatic effects: Glu335Gln causes the toxin to transform into a persistent apoenzyme devoid of zinc, and Tyr350Ala has no hydrolytic activity. The structural analysis of several mutants has led to a better understanding of the catalytic mechanism of this family of proteins. The residues forming the S1′ subsite have been identified by comparing this structure with a thermolysin−inhibitor complex structure.

Clostridial neurotoxins comprising seven antigenically different botulinum neurotoxins (BoNTs)[1] and the tetanus neurotoxin (TeNT) are unique zinc endopeptidases. They all share significant sequence homology and structural and functional similarities and cleave one of the three proteins of the SNARE complex required for docking and fusion with target membranes for neurotransmitter release (*1*). Cleavage of any one of the SNARE proteins inhibits the exocytosis of neurotransmitters causing paralysis (*2*). Neurotoxins are unique since they cleave specific substrates precisely at a particular peptide bond. What confers this unique property is still not understood which is essential if a common therapeutic agent is to be developed for all of them since they are considered potential biowarfare threats. The active sites of the catalytic domains of at least four of them have been analyzed crystallographically (*3*−*7*).

Various serotypes of botulinum neurotoxins share nearly 32% sequence identity and 53% similarity (*8*, *9*). A comparison of the catalytic domains of botulinum neurotoxins for which experimental structures are available shows a deep cavity through which substrates could access the active site. Despite high sequence and structural similarity each serotype has a specific substrate and cleavage site selectivity. This suggests that the enzyme has unique recognition site(s) that help(s) in the substrate binding and docking leading to ultimate cleavage of the scissile bond. The recognition site is presumably the SSR sequence (SNARE secondary recognition) motifs present in each neuronal synaptic vesicle fusion protein (*10*). The specificity of the botulinum neurotoxins toward the SSR sequence is very low, and almost each one is able to bind to all SSR sequences irrespective of substrate specificity (*10*). However, each serotype requires a minimum length of the substrate which includes a SSR sequence for the optimum activity (*11*, *12*). The clostridial neurotoxin substrate specificity could also be dictated by the complementarity between the substrate and the enzyme.

The active site of botulinum neurotoxins is comprised of the HExxH+E zinc-binding motif. In the case of BoNT/E-LC His211, His215, and Glu250 directly coordinate with zinc, and a nucleophilic water hydrogen bonded to Glu212 acts as the fourth ligand (*3*). The general conformation and the active site residues are conserved in all clostridial neurotoxins.

Mutational studies have been carried out extensively on BoNT/A and TeNT-LC (*13*−*16*). The effect of mutating residues Arg362, Tyr365, and Glu350 in BoNT/A on the proteolytic activity has been studied (*13*, *14*). It was also established that Glu224Gln completely abolishes the activity in BoNT/A; Glu224, though not directly coordinated to zinc, is hydrogen bonded to the nucleophilic water coordinated to zinc (*17*). Mutational studies on thermolysin, a metallo-

† Research supported by the U.S. Army Medical Research Acquisition Activity (Award No. DAMD17-02-2-0011) under DOE Prime Contract No. DE-AC02-98CH10886 with Brookhaven National Laboratory. T.B. was supported by Grant RGY0027/2001 from the Human Frontier Science Program.

* To whom correspondence should be addressed. E-mail: swami@bnl.gov. Telephone: (631) 344-3187. Fax: (631) 344−3407.

‡ Brookhaven National Laboratory.

§ Medizinische Hochschule Hannover.

[1] Abbreviations: BoNT, botulinum neurotoxin; TeNT, tetanus neurotoxin; SNARE, soluble NSF attachment protein receptor; SSR, SNARE secondary recognition; LC, light chain; SNAP-25, 25 kDa synaptosome-associated protein; PMSF, phenylmethanesulfonyl fluoride; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid.

10.1021/bi050253a CCC: $30.25
Published on Web 05/19/2005

Table 1: Complementary Primer Sequences Designed and Used for Generating the Different Mutations in BoNT/E-LC

mutation	primer sequence
Glu249Ala	F: 5′-CAAATATAAGAGGTACAAATATTGCAGAGTTCTTAACTTTTGGAG-3′
	R: 5′-CTCCAAAAGTTAAGAACTCTGCAATATTTGTACCTCTTATATTTG-3′
Tyr350Ala	F: 5′-GTTAAATGTAGGCAAACTGCCATTGGACAGTATAAATACTTCAAACTTTC-3′
	R: 5′-GAAAGTTTGAAGTATTTATACTGTCCAATGGCAGTTTGCCTACATTTAAC-3′
Arg347Ala	F: 5′-CTAAATTTCAAGTTAAATGTGCGCAAACTTATATTGGACAG-3′
	R: 5′-CTGTCCAATATAAGTTTGCGCACATTTAACTTGAAATTTAG-3′
Glu335Ala	F: 5′-ATTATACAGCTTTACGGCATTCGATTTAGCAACTAAATTT-3′
	R: 5′-AAATTTAGTTGCTAAATCGAATGCCGTAAAGCTGTATAAT-3′
Glu335Gln	F: 5′-ATTATACAGCTTTACGCAATTTGATTTAGCAACTAAATTT-3′
	R: 5′-AAATTTAGTTGCTAAATCAAATTGCGTAAAGCTGTATAAT-3′
Glu158Ala/Thr159Ala/Asn160Ala	F: 5′-TAGAGAAATATTGGAACTGGCAGCTGCAAATAAATCAGGCTCTGC-3′
	R: 5′-GCAGAGCCTGATTTATTTGCAGCTGCCAGTTCCAATATTTCTCTA-3′

protease whose active site geometry is very similar to botulinum neurotoxins and neprilysin, a related protein, have also been reported. Both of them contain the same kind of zinc-binding motif. Residues His231 and His711 of thermolysin and neprilysin play a crucial role in substrate stabilization (*18*−*21*). In addition, there are other residues which may have corresponding residues in BoNTs. The catalytic mechanism of these two proteins and botulinum neurotoxins appears to be similar with some differences.

A structural comparison of at least three serotypes reveals several conserved residues at/near the active site of the enzymes (*3*) and also at distances >10 Å from the active site zinc. Since the catalytic activity and the kinetic parameters of an enzyme rely on the charged groups present in the active site stabilizing the transition state as acid or base catalysts in the reaction, it is possible to modify the catalytic activity by changing the properties of the active site residues without structural alteration. The proteolysis also depends on the intricate electrostatic potential developed by short-range effects such as hydrogen bonds or hydrophobic interactions, as well as long-range interactions. It is difficult to explain fully the change in the electrostatic interactions due to change in amino acid residues in a protein in an electrostatic model without accounting for all possible parameters contributing to dynamics of the protein (*22*). Accordingly, a major part of the enzyme may be involved in optimizing the catalytic center.

A mutational study on Glu212 has been reported for BoNT/E-LC (*3*). In the same study, several residues in the active site and their interactions with other residues that will be common in all clostridial neurotoxins were pointed out. This report focuses mainly on mutation of some of these residues. These are Glu158Ala/Thr159Ala/Asn160Ala (a triple mutant), Glu249Ala, Glu335Ala, Glu335Gln, Glu212Gln/Glu335Gln (a double mutant), Arg347Ala, and Tyr350Ala. We have analyzed the functional role of these residues systematically by mutating one residue at a time and in combination and by determining the three-dimensional structures of five mutants (Glu158Ala/Thr159Ala/Asn160Ala, Glu335Ala, Glu335Gln, Arg347Ala, and Tyr350Ala) to understand the conformational or other environmental changes in the active site to establish structure−function correlation. Since these residues are mostly conserved, we can define their roles in all serotypes by extension.

EXPERIMENTAL PROCEDURES

Site-Directed Mutagenesis. The complementary PCR primers (Invitrogen) were designed for the missense mutations Glu158Ala/Thr159Ala/Asn160Ala, Glu249Ala, Glu335Ala, Glu335Gln, Arg347Ala, and Tyr350Ala (Table 1). The native pET-9c-LC (*23*) was used as template along with the Quick-Change site-directed mutagenesis kit (Stratagene, La Jolla, CA) for mutagenesis according to the manufacturer's manual. The double mutant Glu212Gln/Glu335Gln was generated with Glu212Gln mutant DNA template and the primers of Glu335Gln mutation. To confirm the mutations, the entire region encoding the mutated light chain (LC) was sequenced in both of the strands using the Big Dye terminator cycle sequencing (Applied Biosystems) method.

Expression and Purification of the BoNT/E-LC Mutant Proteins. The plasmid DNAs were transfected into *Escherichia coli* BL21(DE3) bacteria for the expression of the protein. The growth conditions of the culture for expression and the method of purification of the protein are as for the wild-type protein (*23*).

The 2×YT medium (1.6% bactotryptone, 1.0% bacto-yeast extract, and 0.5% NaCl) was used to grow the cells at 37 °C with shaking until the A_{600} reached 0.6. IPTG (1 mM) was added at this point, and cells were incubated at 20 °C for an additional 12 h. The induced cells were pelleted down by centrifugation at 5000 rpm at 4 °C for 10 min. The cell pellet was suspended in 20 mL of lysis buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 5 mM benzamidine, 0.5 mM PMSF, 1 μg/mL pepstatin A) supplemented with two tablets of protease inhibitor cocktail (Roche), 2 mL of Bugbuster (Novagen), 6 mM iodoacetamide, and 2 μL of Benzonase (Novagen) and incubated at room temperature for 20 min. The lysate was centrifuged at 16500 rpm for 30 min to remove insoluble cell debris. The protein was purified using a Ni-NTA agarose (Qiagen) column and elution with increasing concentrations of imidazole (10−250 mM). Further purification of wild type and mutants was done by size-exclusion chromatography (Superdex-75) employing the Akta-Prime purifier system (Amersham). The buffer was also exchanged from phosphate to HEPES buffer (20 mM HEPES, pH 7.2, 200 mM NaCl) in this process. Proteins were stored at −70 °C until used. A high-level expression of protein has been achieved (>15 mg/500 mL of culture).

Construction of N- and C-Terminal 6×-His-Tagged SNAP-25. The pET28-b vector which has a 6×-His tag at both N- and C-terminal ends has been used. The PCR amplification of the SNAP-25 gene using the pGEX-2T-SNAP-25 vector (kindly provided by Dr. Eric Johnson, University of Wisconsin) as template was performed using the following primers: forward, 5′ CCG CGT **GCT AGC** ATG GCC GAG G 3′, and reverse, 5′ GGC A**CT CGA G**AC CAC

Table 2: Data Collection and Refinement Statistics

	E335Q	E335A[a]	R347A[a]	triple mutant	Y350A
crystal data					
a, *b*, *c* (Å)[b]	88.8(1), 144.6(2), 83.31(1)				
space group	$P2_12_12$				
PDB ID code	1ZL5	1ZN3	1ZKW	1ZKX	1ZL6
resolution (Å)	50−2.6	50−2.6	50−1.93	50−2.5	50−2.4
no. of reflections	30978	34510	72275	36986	42956
I/σ[c]	5.8 (2)	9.0 (3)	6.4 (2)	5.4 (2)	9.0 (3)
R_{merge}[d]	0.12 (0.43)	0.11 (0.50)	0.09 (0.50)	0.07 (0.42)	0.08 (0.43)
redundancy	6.9	6.6	5.0	7.3	5.9
refinement statistics					
resolution (Å)	50−2.6	50−2.6	50−2.2	50−2.5	50−2.4
no. of reflections	28511	32370	49723	26893	38653
completeness (%)	84.3 (55.3)	95.7 (95)	96 (82.4)	74 (54.7)	90 (89)
R-factor[e]	0.20	0.22	0.23	0.20	0.23
R_{free}	0.27	0.26	0.28	0.25	0.28
no. of protein atoms	6563	6463	6571	6545	6305
no. of heteroatoms	9	7	9	6	3
no. of water molecules	250	350	263	244	207
RMSD bonds	0.007	0.008	0.007	0.007	0.007
RMSD angles	1.34	1.40	1.29	1.33	1.29
Ramachandran plot					
most favored (%)	85	85	85	84	85
additionally allowed (%)	14	14	14	15	15

[a] Outermost shell reflections not included for refinement. [b] Mean values for the five structures are given. [c] Values within parentheses are for the outermost shell. [d] $R_{merge} = \sum_j(|I_h - \langle I\rangle_h|)/\sum I_h$, where $\langle I_h\rangle$ is the mean intensity of reflection *h*. [e] *R*-factor $= \sum|F_o - F_c|/\sum|F_o|$.

TTC CCA GC 3′ (the underlined sequence represents *Nhe*I and *Xho*I restriction sites). The PCR sample and the vector were digested with the same restriction enzymes and were ligated together using the T4-ligation system (New England Biolabs). The clone was confirmed for the correct sequence by DNA sequencing of both of the strands by the Big Dye terminator cycle sequencing (Applied Biosystems) method.

Expression and Purification of SNAP-25. The pET28-b-SNAP-25 clone was transformed to *E. coli* BL21(DE3) cells and expressed. The purification protocol is the same as for BoNT/E-LC (*23*). The protein is stable for more than 6 months when stored between −80 and −20 °C. On a native gel the protein showed a band corresponding to ∼52 kDa, suggesting it to be a dimer in solution. On SDS−PAGE gel it shows the expected ∼26 kDa size. The total recovery of protein was found to be ∼20 mg/L of culture.

Enzymatic Activity of Wild-Type vs Mutant Proteins. The proteolytic activity of the wild type and mutants of BoNT/E-LC was assayed in vitro on its substrate SNAP-25. The assay was performed in a final volume of 20 μL of activity assay buffer (20 mM HEPES, pH 7.4, 2 mM DTT) containing 2 nM concentration of LC and 7 μM concentration of SNAP-25 at 37 °C for 60 min. Trials were carried out with and without zinc acetate in the assay buffer. Since the difference was marginal or none, the results reported here are with assay buffer devoid of zinc acetate. However, the mutant protein concentration in the reaction mixture was varied in another experiment from 5 to 500 nM, since at 2−5 nM some of the mutants did not exhibit any hydrolytic activity. Here, the cleavage of SNAP-25 by the LCs was determined by incubating the samples at 37 °C for 20 min. The reactions were stopped by adding 10 μL of 3× concentrated SDS−PAGE sample buffer containing 1 mM EDTA to chelate the Zn cofactor, if present. The extent of cleavage was then evaluated following electrophoresis on 4−20% Tris−glycine SDS−PAGE gels (Cambrex) by the appearance and intensity of a new band at ∼23 kDa due to the cleavage of SNAP-25 between aa 180 and 181. The cleaved and uncleaved substrate fractions were quantified using densitometry, and the kinetic parameters were derived using Lineweaver−Burk plots.

Binding Assay of Mutants. To ascertain that the mutant's reduced or loss of catalytic activity is not due to their inability to bind to the substrate, a binding assay was performed. The wild-type enzyme and the mutants were mixed in 1:1 ratio and then incubated with SNAP-25 in 1:1000 ratio in the activity assay buffer for 15 min at 37 °C. As a control, the same experiment was repeated with the wild-type enzyme premixed with an unrelated protein BSA or lysozyme. All of the reactions were stopped and run on SDS−PAGE gels to compare the results.

Crystallization and Data Collection of Glu335Gln, Glu335Ala, Arg347Ala, Tyr350Ala, and Glu158Ala/Thr159Ala/Asn160Ala Mutants. Crystallization screening was carried out by the sitting drop vapor diffusion method using Hampton Research crystallization screens. Diffraction quality crystals were obtained at room temperature using 0.5 M ammonium sulfate, 1.0 M Li_2SO_4 and 0.1 M sodium citrate trihydrate at pH 5.6 as precipitant, and crystals grew to their full size in 2−6 days. Crystals belong to the space group $P2_12_12$, same as wild type (*23*). The cell parameters for crystals of mutant proteins and wild-type LC are similar.

Diffraction data from crystals of mutants were collected at the NSLS beamlines. Crystals were briefly transferred to the mother liquor containing 20% glycerol and were mounted on a nylon loop and flash frozen immediately by plunging into liquid nitrogen. Data covering at least 180° rotation in ϕ were collected, for an oscillation range of 1° per frame at $\lambda = 1.1$ Å. Data were processed, scaled, and merged with HKL2000 (*24*). Details of data collection statistics are given in Table 2.

Structure Determination of Mutant Proteins. Since the mutant crystals were isomorphous to the wild type, the BoNT/E-LC model (PDB ID 1T3A) with appropriate muta-

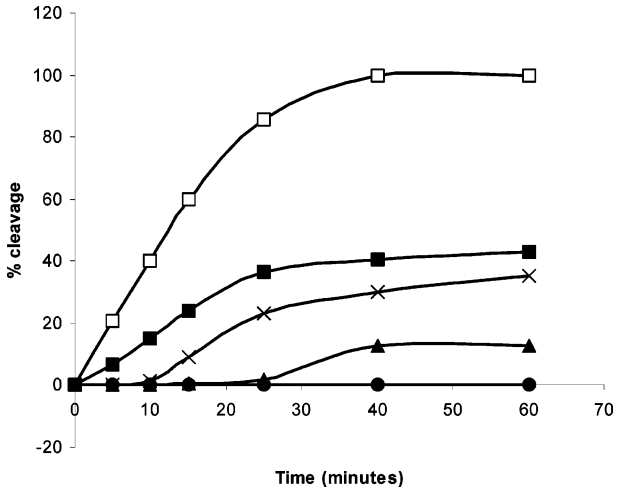

FIGURE 1: Catalytic activity of wild-type BoNT/E-LC and various mutants. The cleavage values were determined in a total volume of 20 μL at 37 °C using 7 μM N- and C-terminal 6×-His-tagged SNAP-25 with 2 nM light chain concentrations. At indicated time intervals aliquots were removed and analyzed on SDS−PAGE gels and with Coomassie blue staining. The quantity of protein was determined using densitometry analysis. The data represent the average of three independent experiments: wild type (□), triple mutant (■), Glu249Ala (×), Glu335Ala (▲) and Arg347Ala, Tyr350Ala, Glu335Gln, and Glu212Gln/Glu335Gln double mutant (●).

tion was used as the starting model. After initial rigid body refinement, the model was refined by simulated annealing using CNS (*25*). Solvent molecules were added also using CNS. Composite omit maps were computed and used to fine-tune the model without model bias using the program O (*26*). The refinement statistics are included in Table 2.

RESULTS AND DISCUSSION

Enzymatic Activity of Wild-Type and Various Mutant Proteins. The hydrolytic activities of the wild-type and mutant proteins at the same concentration are shown in Figure 1. Triple mutant and Glu249Ala showed ~43% and 36% activity as compared to wild type, whereas Tyr350Ala, Arg347Ala, and Glu335Gln failed to show any activity in that particular experiment. Glu335Ala showed ~13% activity with respect to wild type.

In another experiment, various concentrations of the mutant proteins (5−500 nM) were used on 7 μM substrate with a fixed time interval (20 min) at 37 °C (data not shown). Wild type could digest the substrate completely at 5 nM enzyme concentration, while Tyr350Ala did not show any activity even at a 500 nM level. For Glu335Gln and Arg347Ala mutants there was no activity up to 50 nM, but residual activity started appearing at 500 nM enzyme concentration with about <5% and 20% of the substrate cleaved, respectively. While the Glu335Ala mutant started showing activity at 50 nM, other mutants such as Glu249Ala and the triple mutant had detectable activity even at 5 nM concentration.

Kinetic Studies for the Wild-Type and Various Mutant Proteins. The enzyme kinetic values were determined by the cleavage rate of SNAP-25 at different concentrations. The derived K_m and K_{cat} values are shown in Table 3. The values for the wild type are slightly different from what we had reported earlier (*23*). This may be because we had used

Table 3: Enzyme Kinetic Parameters of the Wild-Type BoNT/E-LC and Various Mutants

light chains	K_m (μM)	K_{cat} (min^{-1})	K_{cat}/K_m (μM^{-1} min^{-1})
wild type	7.9 ± 0.4	257.11 ± 21	32.75
Glu335Gln	10 ± 0.5	0.045 ± 0.003	0.0045
Glu335Ala	9.7 ± 0.8	8.66 ± 0.2	0.89
Tyr350Ala	not detectable	not detectable	not detectable
Arg347Ala	11.2 ± 0.2	0.35 ± 0.045	0.031
triple mutant	11.3 ± 1.0	50.17 ±	4.43
Glu249Ala	11.2 ± 2.0	24.17 ± 5.9	1.62
Glu212Ala/ Glu335Gln	not detectable	not detectable	not detectable

SNAP-25 with a GST tag earlier while we are using SNAP-25 with an N- and C-terminal 6×-His tag now. For mutants the assay was repeated by adding (an increasing amount of) zinc acetate (50−100 μM) to the reaction buffer, but this had no effect on the activity.

Two of the mutations with drastic effect on the activity are Tyr350Ala, which failed to show any activity, and Glu335Gln, with ~7000-fold less K_{cat}/K_m than wild type. However, substitution of Glu335 to Ala reduced the K_{cat}/K_m by ~40-fold only. The K_m values for all of the mutants were only marginally higher than that for the wild type, suggesting similar affinity to the substrate. Any change in activity was due to changes in the K_{cat} values. However, K_m for Tyr350Ala could not be determined. Another severe phenotype was exhibited by the Arg347Ala mutant, which has ~1000-fold lower catalytic efficiency than the wild type. On the basis of severity phenotype the mutants in this study can be ordered as Glu335Gln/Glu212Gln and Tyr350Ala > Glu335Gln > Arg347Ala > Glu335Ala > Glu249Ala > triple mutant.

Substrate Binding Assay for Mutant Proteins. The substrate binding assay for the various mutants such as Arg347Ala, Tyr350Ala, Glu335Gln, and Glu212Gln/Glu335Gln showed that they compete for binding to the substrate with the wild-type enzyme (data not shown). This is also supported by the K_m values of the hydrolytically active (at least to some extent) mutants, which are similar to the wild-type value.

Structural Description of Various Mutants. In general, there is no global conformational change in any of the mutants compared to the wild-type structure (Figure 1 in ref *3*) except for some side chain orientation and the position of nucleophilic water at the active site. A few extra residues appear in the loop regions for some mutant structures, which is not uncommon. Description and discussion of each structure follows.

(A) Arg347Ala. Arg347 was selected for the study since this residue has been implicated in catalytic activity both by a mutagenesis study (BoNT/A) and by similarity to thermolysin Arg203 (*13*, *18*−*21*, *27*). In BoNT, Arg347 is in the secondary coordination sphere of zinc. Glu335 acts as a bridge between Arg347 and His211 and is hydrogen bonded to them, stabilizing the active site (*3*). It is possible that any perturbation in this interaction might affect the active site architecture and, consequently, the activity. However, the terminal nitrogen atoms of the guanidino group of Arg347 are more than 7 Å away from either zinc or nucleophilic water in BoNTs. The active site geometry is almost preserved in the Arg347Ala mutant structure. The coordination of zinc remains the same except that the nucleophilic water has

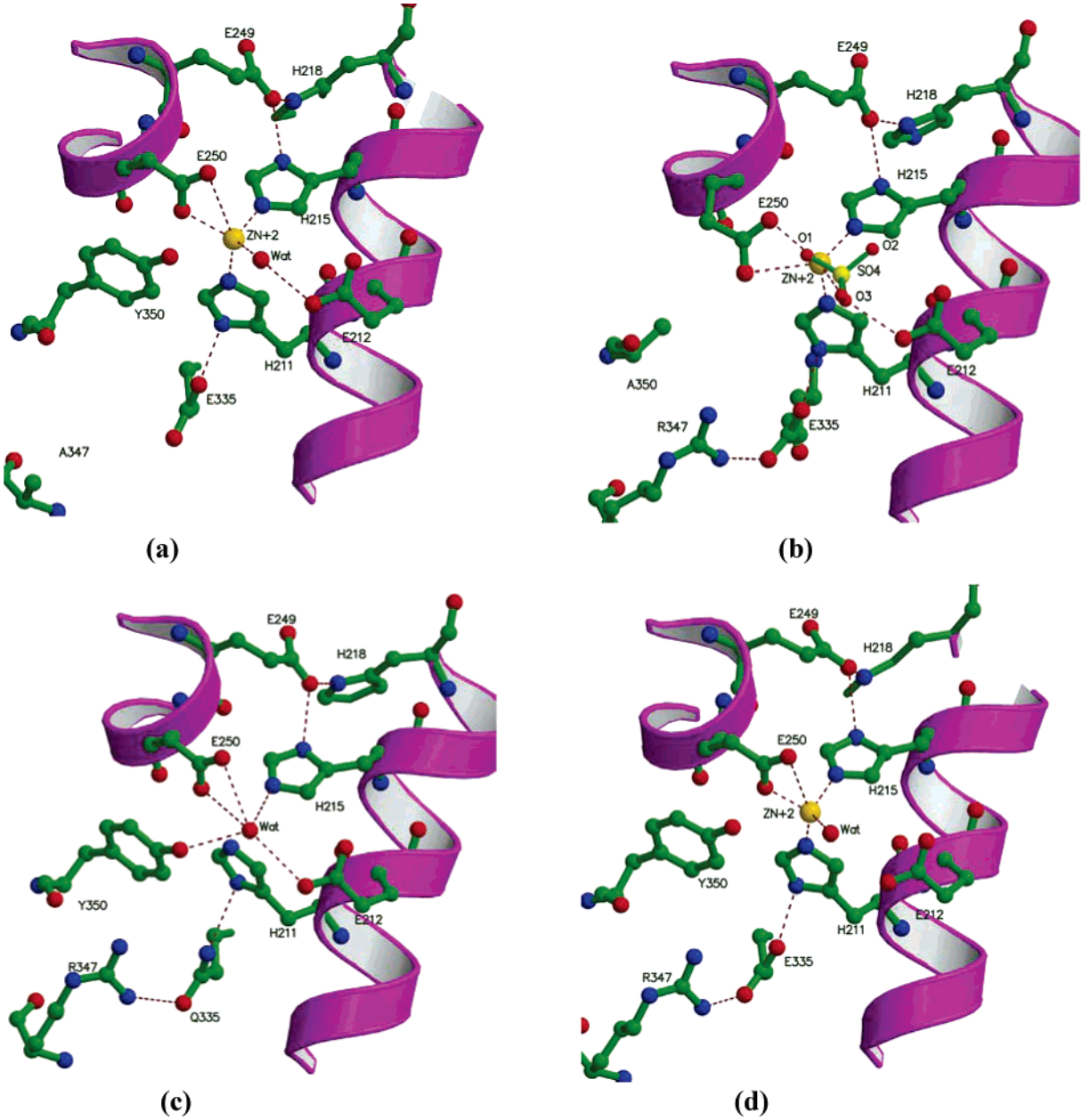

FIGURE 2: Active site of the four mutants: (a) Arg347Ala, (b) Tyr350Ala, (c) Glu335Gln, and (d) Glu158Ala/Thr159Ala/Asn160Ala triple mutant. Carbon, nitrogen, and oxygen atoms are shown as spheres in green, dark blue, and red. Zinc is in yellow, and dashed lines represent coordination and hydrogen bond contacts. For Tyr350Ala, the active site of molecule B is given. All distances are given in Table 4. These figures were generated with MOLSCRIPT and Raster3D (*34*).

moved away to 2.6 Å, changing from coordination to hydrogen bonding distance (Figure 2a). It was expected that the loss of hydrogen-bonding interaction between Arg347 and Glu335 would affect the hydrogen bond contact between Glu335 and His211. However, this hydrogen bond is intact, and the orientation of His211 is unaltered and similar to that of the wild-type structure (*3*).

(B) Tyr350Ala. Mutation of the residue equivalent to Tyr350 has been studied both in BoNT/A and in TeNT (*13*−*15*). When it was substituted with Phe, there was 25−30% residual activity in both BoNT/A and TeNT. Mutation with Ala had different effects in BoNT/A and TeNT. While the activity was completely lost in TeNT (*15*), about 20−30% was retained in BoNT/A (*14*). In BoNT/E-LC the effect is similar to that of TeNT, and the activity is barely detectable even at high concentration (5 μM) of the mutant enzyme (Tyr350Ala).

As in the case of Arg347Ala, the conformation of the active site remains almost the same. The coordination distance between zinc and histidines is not changed. However, as a consequence of the loss of the phenolic ring and hence the aromatic−anionic interaction stabilizing the position of Glu250, the side chain of Glu250 takes a different rotamer position and moves closer to the original position of the Tyr350 side chain. It also increases the distance between zinc and the carboxylate group of Glu250 (molecule A). Interestingly, in the crystal structure of Tyr350Ala a sulfate ion replaces the nucleophilic water in molecule A similar to that observed in the BoNT/B holotoxin structure (*28*). The sulfate ion may have been incorporated during crystallization, since ammonium and lithium sulfates were present in the crystallization condition (Figure 2b).

The nucleophilic water is intact in molecule B of the LC dimer, though displaced to 3.21 Å from zinc. It moves closer

to Glu212 and is at a distance of 2.42 Å, making a very strong hydrogen bond. This is similar to movement of the water during the transition state, allowing the scissile carbonyl carbon to form a tetrahedral geometry (*29*). This movement of water closer to Glu212 is because of the loss of interaction between Tyr350 OH and the nucleophilic water that was present in the wild-type structure. This may be one reason a sulfate ion is able to displace the water in molecule A of Tyr350Ala unlike other BoNT/E structures where the water is tightly kept in position by tyrosine even though the crystallization conditions are the same. The packing consideration could have prevented the replacement of water in molecule B.

(C) Glu335Ala. Residues Glu249 and Glu335 are in the secondary coordination sphere in clostridial neurotoxins. Glu249 helps in positioning and orienting His215 for coordination with zinc, while Glu335 does the same for His211. In addition, these two are charged amino acids and are about 8 Å away from the zinc. As in the case of Glu212, changing them to nonconservative residues might alter the electrostatic potential near the zinc atom affecting the catalytic activity significantly. The Glu249Ala mutation caused the activity to decrease 15-fold. Glu249 is a conserved residue in all BoNTs except in serotype C. In the case of serotype C this residue is an alanine, and not surprisingly BoNT/C is less active than other BoNTs (*11*). However, the structure of Glu249Ala was not determined.

The crystal structure of Glu335Ala is similar to the wild type except for some minor changes in the side chain conformation of the active site residues. As in the case of Arg347Ala, the nucleophilic water has moved away from the active site zinc closer to Glu212. Active site zinc is intact, and the orientation of His211 is unaltered. A strong residual density persisted in the difference Fourier map and was modeled as chlorine because of its location. Interestingly, this chlorine ion is bridging the side chains of His211 and Arg347. In molecule B, this chlorine ion may be disordered as also the nucleophilic water.

(D) Glu335Gln. Interestingly, when Gln replaces Glu (Glu335Gln) there was a drastic drop in catalytic efficiency (~7000-fold). Though the active site geometry remained the same, it was devoid of zinc (Figure 2c). Neither the composite omit map nor the difference Fourier computed with the final model showed any electron density at the expected zinc position (Figure 3). Though we did not determine the zinc content by spectroscopic experiments, we routinely scan the crystal (in this case even the protein at 10 mg/mL) near the absorption edge of zinc at the synchrotron beamline during X-ray diffraction data collection. While wild type and other mutants gave a good signal near the zinc absorption edge (9659 eV), Glu335Gln and Glu212Gln/Glu335Gln were noisy with no detectable signal. Glu335 makes hydrogen bonds with Arg347 and His211 in the wild-type structure. Though the type of contacts may be different, Gln335 also makes hydrogen bonds with both Arg347 and His211 (significantly longer than wild type, Table 4), but the orientation of His211 has changed (χ_2 by 20°) and probably its protonation state also. It is intriguing that the nucleophilic water is still present, though moved away from the original position by about 0.8 Å (Figure 4a). It makes strong hydrogen bonds with His215, Glu250, and Glu212 and also with Tyr350 (Table 4). This suggests that the nucleophilic water is compensating for the loss of the zinc ion by interacting more closely with these residues. This is similar to the apo BoNT/B structure, though in apo BoNT/B, Glu267 (corresponding to Glu250 of BoNT/E) had taken a different rotamer position from the holotoxin moving closer to Gln264 (*30*). In BoNT/E-LC, the position and orientation of Glu250 are unaltered, but both of the carboxylate oxygens make strong hydrogen bonds with the nucleophilic water. A similar water molecule has been observed in the apo BoNT/B crystal structure also.

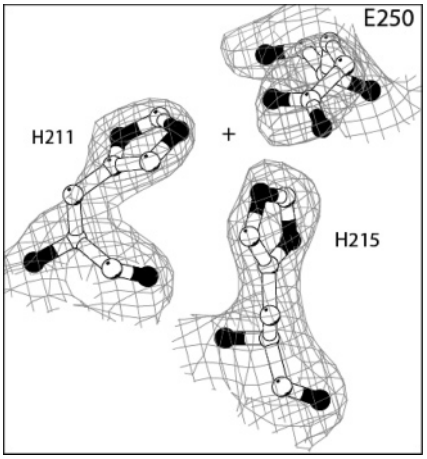

FIGURE 3: Composite omit map for Glu335Gln. Map contoured at 1σ. Zinc position corresponding to the wild-type structure is shown with a + sign for reference. There is no electron density corresponding to this in the Glu335Gln structure.

Table 4: Comparison of Interactions (Å) between Crucial Residues at the Active Site[a]

as in wild type	wild type	R347A	E335A	E335Q	triple mutant	Y350A (B)
Zn−H211 NE2	2.18	2.30	2.36		2.16	2.06
Zn−H215 NE2	2.16	2.06	2.26		2.26	2.11
Zn−E250 OE1	2.22	2.24	3.79		2.64	2.73
Zn−E250 OE2	2.46	3.13	2.20		2.06	2.40
Zn−Nu wat	2.17	2.60	2.47		2.81	3.21
Nu wat−E212 OE1	4.03	3.79	3.83	3.79	3.97	3.79
Nu wat−E212 OE2	2.86	2.96	3.06	3.04	3.32	2.42
Nu wat−Y350 OH	3.58	3.53	4.44	2.90	3.92	
Nu wat−E250 OE1	3.15	3.86	3.84	3.00	3.58	4.47
Nu wat−R347 NH2	7.06		7.01	7.16	7.70	7.80
Nu wat−E250 OE2	3.66	4.02	3.86	3.19	4.14	5.32
Nu wat−H215 NE2	3.74	4.04	4.35	3.29	4.41	4.54
E335OE1−H211 ND1	2.61	2.52		2.99	2.59	2.59
E335 OE2−R347 NH1	2.99			3.39	2.94	3.13
E249 OE1−H215 ND1	2.76	2.79	2.93	2.85	3.01	3.10
E249 OE1−H218 ND1	2.73	2.69	2.73	2.80	2.97	2.82
Y350 OH−Zn	3.94	3.99	4.87		4.05	
Y350 OH−E250 OE1	3.15	3.38	3.66	3.48	3.69	
Y350 OH−E250 OE2	3.43	3.76	3.63	3.80	3.53	
R347 NH2−Zn	7.08		7.07		7.19	7.01

[a] The two molecules in the asymmetric unit are identical within experimental error (RMSD = 0.8 Å), and accordingly, the interactions for monomer A are given for all structures except for Y350A. The nucleophilic water is displaced by a sulfate ion at the active site of monomer A in Y350A while B retains the water. Thus, for Y350A, interactions in molecule B are given. Nu wat is nucleophilic water.

(E) Glu158Ala/Thr159Ala/Asn160Ala, a Triple Mutant. These residues were considered on the basis of our modeling of the transition state in BoNT/B and the thermolysin structure (*29*). The active sites of thermolysin and BoNT/E-LC are similar though with some differences. Asn112 and Ala113 of thermolysin, which make contact with the scissile

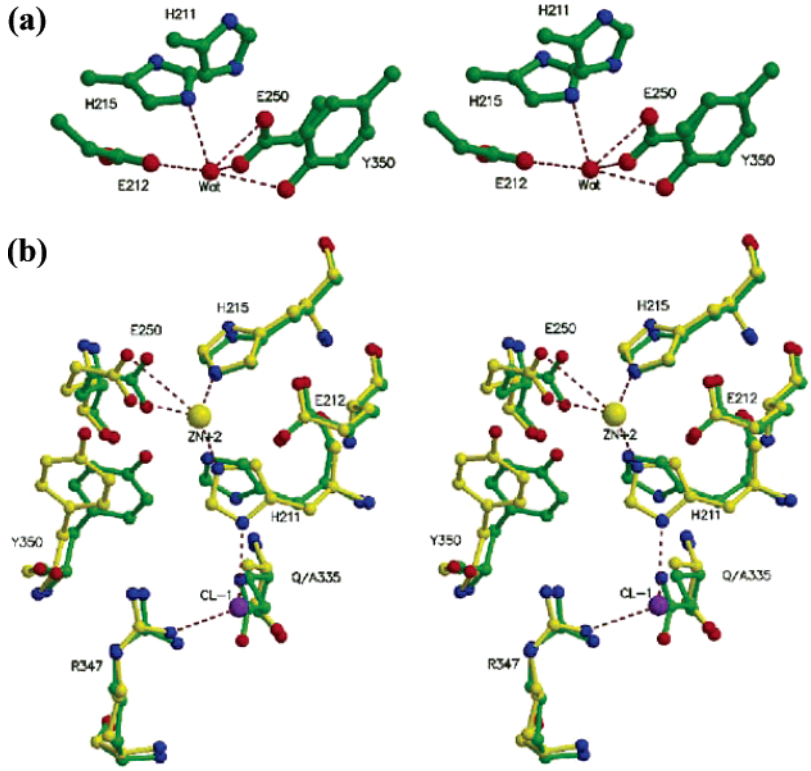

FIGURE 4: (a) Stereoview of the active site residues in Glu335Gln. The active site zinc is absent in this structure. The nucleophilic water has moved and makes strong hydrogen bonds with His211, Glu212, Glu250, and Tyr350. (b) Stereoview of the superposition of Glu335Ala (yellow) and Glu335Gln (green). The zinc ion corresponding to Glu335Ala is also shown in yellow and the chlorine ion in purple. Nucleophilic water molecules of the two structures are not shown for clarity. The change in orientation of His211 is apparent in the figure.

bond nitrogen, are positioned similarly to Glu158 and Thr159 in BoNT/E. Moreover, Glu158, a charged group, is <10 Å from the catalytic center. On the basis of these facts the triple mutant was constructed. The catalytic activity decreased by ~7-fold only. The crystal structure showed that the conformation of the active site is maintained but the nucleophilic water moved to hydrogen-bonding distance from zinc (2.81 Å) (Figure 2d). In the case of the triple mutant, since the hydrogen bond with the scissile peptide bond nitrogen is from the main chain O of Thr159, the activity decreases because of the loss of surface charge on Glu158Ala.

Implications on the Catalytic Mechanism. (A) Arg347 and Tyr350 Stabilize the Substrate Binding. In Arg347Ala, except for the movement of nucleophilic water, there is no major change in the active site conformation though the activity of the mutant is reduced 1000-fold. The effect is thus 13 times larger than that observed for the corresponding mutation in BoNT/A (*13*). It has been suggested that the arginine residue corresponding to Arg347 could be polarizing the nucleophilic water based on an inhibitor binding in thermolysin (*31*). However, in the present study, the low activity suggests that it may not be playing the role of polarizing the nucleophilic water, which is also supported by similar observations in BoNT/A-LC (*13*). The drastic reduction in activity without an accompanying structural change suggests that it plays an important role in substrate stabilization.

When the LC separates from the holotoxin, the loops rearrange and bring Tyr350 closer to the active site (*4*, *28*). The hydroxyl group of Tyr350 is 3.94 and 3.58 Å from zinc and the nucleophilic water, respectively, in the wild-type BoNT/E-LC (*3*). In addition to the hydroxyl group of Tyr350 being at hydrogen-bonding distance from the nucleophilic water, the aromatic−anionic interaction between Tyr350 and Glu250 helps to stabilize the position of Glu250. In BoNT/A, the reduction in activity of Tyr365Phe may be due to the loss of hydrogen-bonding interaction between the OH group and the nucleophilic water. In the BoNT/E-LC mutant Tyr350Ala, the complete loss of activity is due to the loss of both interactions. The side chain of Glu250 moves considerably, affecting the overall interactions with other active site residues. In TeNT the loss of activity for Tyr375Ala may be due to similar reasons. The reason for residual activity in the BoNT/A mutant Tyr365Ala is not clear (*14*). It may be because of some unidentified interaction compensating for the lost interaction. In our earlier study with Glu212Gln, which completely abolished the activity, we concluded that the proton donor must be Glu212 as in thermolysin (*3*). While mutation of Glu212 in BoNT/E and Glu224 in BoNT/A abolished the activity completely, the result of mutation of Tyr350 depends on the kind of mutation. Since the absence of the hydroxyl group (Tyr350Phe) only reduces the activity, Tyr350 may be providing the stabilizing force rather than donating a proton to the leaving group. This conclusion is also similar to that in thermolysin where the equivalent His231 plays a similar role (*21*).

Catalysis is a three-step process: the substrate recognition, docking to the active site, and the cleavage at a specific peptide bond. The optimal docking will bring about the maximum cleavage. The catalysis will be affected on the basis of the relative contribution of a particular residue in actual docking/stabilization of the substrate to the active center. We have shown that substitution of Tyr350 with Ala abolishes the catalysis, revealing the importance of this

residue in catalysis. The results suggest that the interactions of the carbonyl oxygen to the zinc and the Tyr350 are crucial for transition state stabilization. The hydroxyl group of this residue has been proposed to be the one involved in donating a proton to the leaving amide group (*4*, *15*). But since the change of tyrosine to phenylalanine does not abolish the activity either in BoNT/A or in TeNT, it may not be the proton donor for the leaving group. In thermolysin and neprilysin the direct interactions of similar residues to the substrate have been proposed (*18*−*21*). We have also proposed previously that Glu212 must be directly involved in proton donation based on the mutation Glu212Gln, which abolishes the activity completely (*3*).

In the initial stages of catalysis the reactants have to pass through a transition state to bring about catalysis. The transition state is often the result of strain or distortion of the reactants to form the particular electronic structure needed for the proper collision and product formation. Therefore, we propose that though the absence of Tyr350 may not affect the substrate docking severely (as seen by our competitive binding study), it is unable to provide the energy required to sustain the strain or distorted state needed for the catalysis, and hence the cleavage of the scissile bond is severely affected.

It is also not clear why the K_m remains the same for Tyr350Ala and Arg347Ala if their role is for substrate stabilization. One explanation is that the loss of one interaction is compensated by another interaction (*19*). However, further studies with inhibitor molecules with mutants are needed to understand this more thoroughly.

(B) Glu335Gln Is an Apoenzyme. Despite the absence of zinc, there was some residual cleavage at high concentration (500 nM) of the mutant enzyme. It is not clear how a zinc endopeptidase devoid of zinc could still have some residual activity. One explanation is that the presence of nucleophilic water and its strong hydrogen-bonding interactions could still help in attacking the carbonyl carbon of the scissile bond. Or, Glu212, which is strongly hydrogen bonded to the nucleophilic water, could still act as a base, polarize the nucleophilic water, and attack the carbonyl carbon without the help of the zinc ion. This is similar to the Glu144Ser active site mutant of the *Bacillus cereus* thermolysin-like neutral protease where albeit in the absence of a charged base some residual activity was detected and attributed to the nucleophilic water molecule strongly hydrogen bonded to Ser144, a neutral amino acid (*32*). However, while the zinc ion was still present in that structure, it is absent in Glu335Gln. Another explanation for the residual activity could be the presence of traces of zinc in the assay buffer solutions, but we have shown that adding increasing amounts of zinc to the assay buffer does not change the (in)activity. In view of this Glu335Gln, an enzyme without the cofactor zinc, should be considered an apoenzyme, a persistent one in that it cannot bind zinc anymore. Even though we have shown that Glu212Gln is completely inactive, we constructed a double mutant Glu212Gln/Glu335Gln which will be both inactive and apoenzyme devoid of zinc. As expected, this double mutant had no activity whatsoever.

The extremely low catalytic efficiency (~7000-fold less) shown by the mutation Glu335Gln along with it being an apoenzyme (devoid of zinc) is baffling. The data are also supported by similar biochemical results obtained with BoNT/A-LC. Mutating the corresponding residue in BoNT/A to Ala/Gln led to drastic loss of activity and lowering of zinc content. The lower thermal stability observed in BoNT/A for this mutation may be due to the loss of zinc (*13*). Lowering of thermal stability has also been observed for apo BoNT/A-LC (*33*). When Glu is changed to Gln, the net charge changes by 1. This affects both the attractive and repulsive forces with the neighboring residues. Since the total electrostatic potential at any point is due to the sum total of all these changes, the environment must have changed. The change of interaction between this residue and the acidic residues would have caused this change. This is a classic example of how a subtle change in electrostatic properties (Glu is charged while Gln is neutral) can cause a major change in the environment. When the electrostatic potential surfaces of the wild-type enzyme are compared to those of mutant structures, it was clear that the electrostatic potential near Gln335 has changed considerably (data not shown). It is puzzling that Glu335Ala retains zinc and consequently more activity than Glu335Gln. The change in charge is the same except that the side chain of Gln335 can interact with surrounding residues unlike Ala335, resulting in different charge distribution in this region. The difference in orientation of His211 in Glu335Ala and Glu335Gln is shown in Figure 4b. The orientation of His211 in wild type is closer to Glu335Ala than Glu335Gln. This is probably because of the presence of a chlorine ion in the Glu335Ala mutant in the position of the side chain carboxylate group of Glu335 of wild type. It makes contacts with Arg347 NH2 and His211 ND1 at distances 3.27 and 3.34 Å, respectively. The reduction in activity may be partly because Tyr350 moves away from the active site (Table 4). The change in conformation and protonation of His211 may be a dominant factor for drastic reduction in Glu335Gln.

(C) Position of Nucleophilic Water Is Crucial for Activity. Movement of nucleophilic water is common to all mutant structures where decreased or loss of activity is observed. In Tyr350Ala and Arg347Ala the water position changes, contributing to the loss of activity. A similar observation has been made with Glu212Gln also (*3*). In Glu335Gln, despite the absence of zinc, the nucleophilic water is retained and makes stronger hydrogen bonds with active site residues, probably leading to some residual activity. In the triple mutant and Glu335Ala the movement of nucleophilic water is accompanied by a decrease in catalytic activity.

Structural Comparison of BoNT/E-LC to Thermolysin and Neprilysin. Clostridium botulinum neurotoxins, thermolysin, and neprilysin all belong to a metallopeptidase family with the zinc-binding motif HExxH+E. The distance between the HExxH motif and the second E varies. But it is informative to compare the three-dimensional structures of these three proteins to bring out similarities and differences. Even though the overall structures are different, the active site geometry is very similar in all of them. To compare the active sites of these three proteins, the residues in the zinc-binding motif (His, Glu, and His) were superposed (Figure 5). It has been hypothesized that His231 of thermolysin might correspond to Arg347 of BoNT/E, but the active site superposition shows that His231 is closer to Tyr350 than Arg347, in agreement with a recent analysis (*7*). The OH group of Tyr350 makes a close contact with the nucleophilic water (and also with the active site zinc). A similar contact is observed in

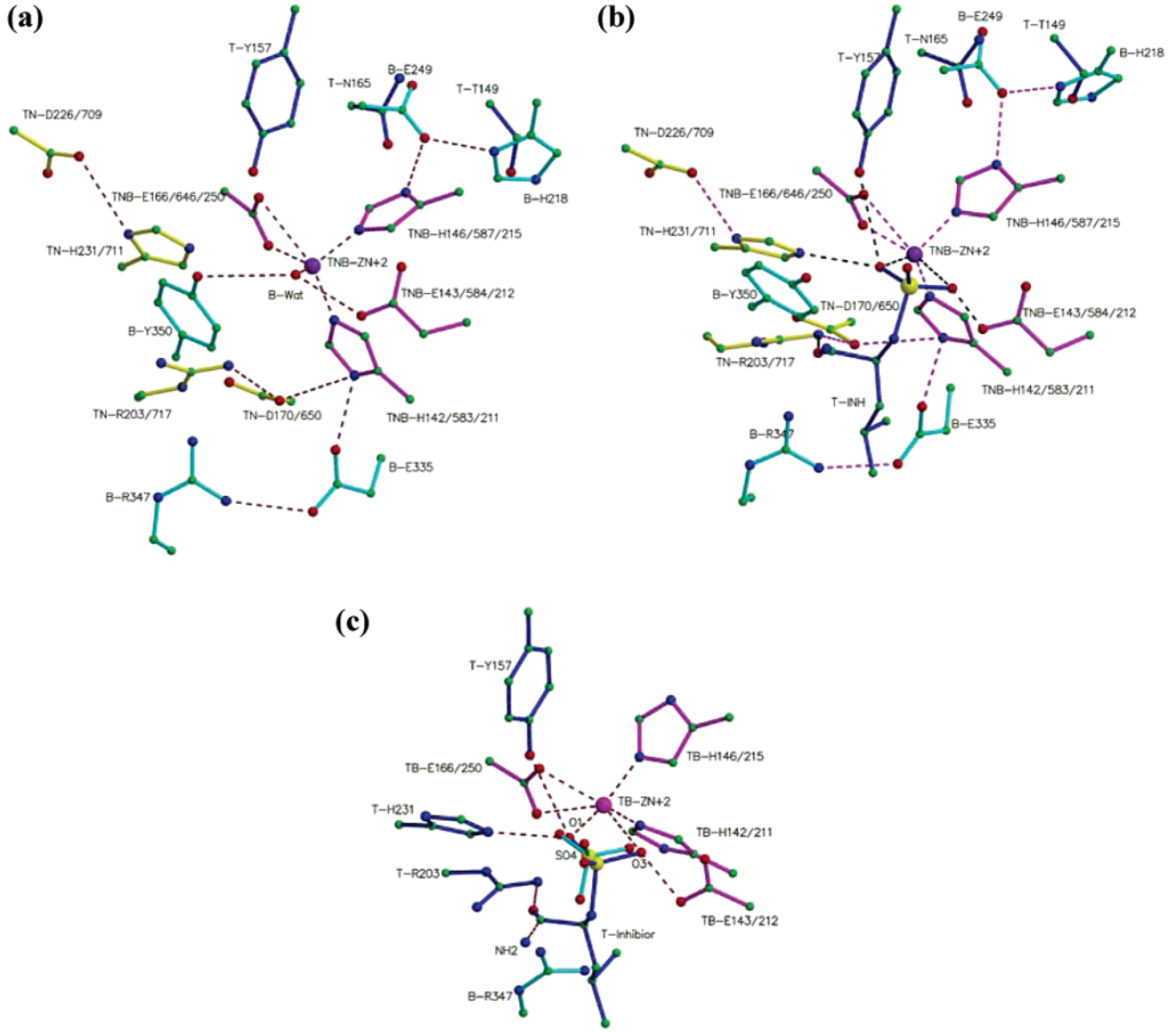

FIGURE 5: Superposition of the active sites of BoNT/E-LC, thermolysin, and neprilysin. Residues coordinating with zinc alone are considered for least-squares fit. (a) Residues common to all three proteins are in magenta and labeled TNB (thermolysin, neprilysin, and BoNT). Residues common to thermolysin and neprilysin (TN) are shown in yellow. BoNT (B) residues are in cyan. Tyr157 (T) which has no corresponding residue in the other two is shown in blue. For clarity only one of the common residues is shown, and the main chain atoms are omitted to avoid cluttering. Carbon, nitrogen, and oxygen atoms are in green, blue, and red, respectively. It is apparent that Tyr350 (B) corresponds to His231/His711 (TN). The hydrogen bond network from Arg203/717 (TN) is similar to that from Arg347 (B). Tyr157 (T) seems to be unique. Detailed comparison is given in the text. (b) As in (a) but shown with the transition state analogue inhibitor molecule of thermolysin (T, in blue) to highlight the interactions. This figure was produced with MOLSCRIPT and Raster3D (*34*). (c) Superposition of the active sites of molecule A of Tyr350Ala and thermolysin with an inhibitor. Residues common to thermolysin (T) and BoNT/E (B) are shown in magenta (for clarity only one of the common residues is shown). Residues of T are in dark blue, and residues of B and the sulfate ion are in cyan. Dotted lines represent coordination or hydrogen bond distances. Carbon, nitrogen, and oxygen atoms are in green, blue, and red, respectively. The sulfate ion positioned and oriented similarly to the phosphate group of the inhibitor in the thermolysin structure (PDB ID 2TMN) provides a model for substrate docking and transition state.

thermolysin (His231) and neprilysin (His711). Arg347 takes part in a hydrogen-bonded network connecting Glu335 and His211 in BoNT/E. Arg203 of thermolysin and Arg717 of neprilysin have a similar network with Asp170 and His142 and with Asp650 and His583, respectively. In view of this Arg347 should play a role similar to Arg203 or Arg717 and Glu335 that of Asp170 or Asp650. Indeed, the effect of mutation on Arg (1000-fold reduction) in botulinum is very similar to that observed in thermolysin and neprilysin (*19*). When Asp170 in thermolysin was changed to Ala, the specific activity decreased 220-fold. When Glu335 which serves a similar purpose in botulinum was changed to Ala, the activity decreased by ~40-fold. However, a mutation similar to Glu335Gln was not done for thermolysin. In thermolysin and neprilysin Asp226 and Asp709 provide stabilizing interactions for His231 and His711. The effect of mutating this Asp was not the same for thermolysin and neprilysin. There is no corresponding residue in botulinum. The reduction in activity due to mutating Tyr350 in botulinum depends on the nature of the mutation, suggesting that it needs both the phenolic ring and the hydroxyl group. Also, Tyr157, which makes a close interaction with nucleophilic water in thermolysin, has no counterpart either in botulinum or in neprilysin. Tyr157 in thermolysin may probably provide

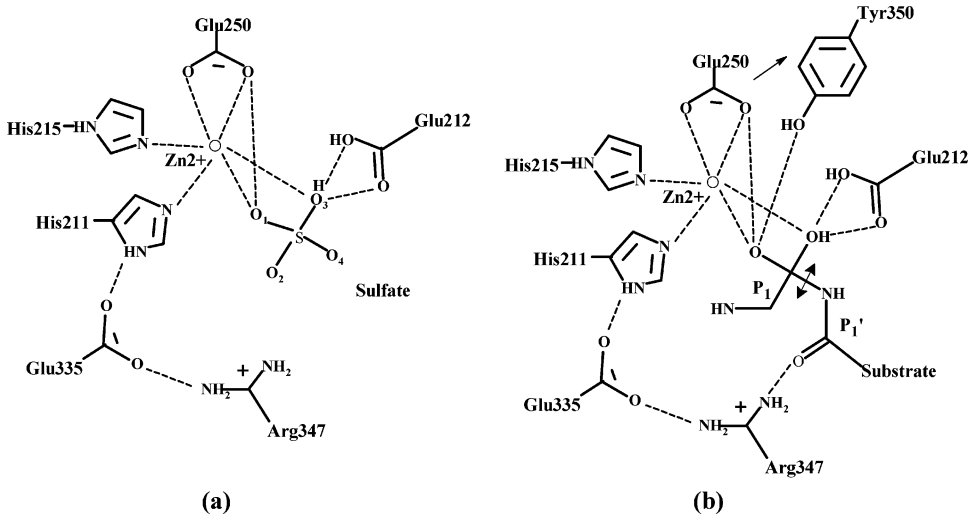

FIGURE 6: Proposed docking of substrate at the active site based on the interactions of the sulfate ion bound to the Tyr350 mutant molecule A and inhibitor-bound thermolysin (2TMN). (a) O1 of the sulfate ion at a distance of 2.4 Å from zinc corresponds to the carbonyl oxygen of the scissile bond (P1) while O3 corresponds to the nucleophilic water which moves closer to Glu212 but is still interacting with zinc. O2 and O4 will correspond to the Cα and scissile bond nitrogen. Thus the sulfate ion mimics the tetrahedral transition state of the substrate. (b) Proposed interactions of the carbonyl oxygens of P1 and P1′ of the substrate during the catalytic pathway. Tyr350 OH interacts with the P1 carbonyl oxygen while the Arg347 NH2 hydrogen bonds with P1′ stabilizing the substrate docking. An arrow mark between Glu250 and Tyr350 represents the anionic−aromatic interaction. The scissile bond is marked with a double-headed arrow mark.

extra stabilizing interactions since the effect of mutating this residue was considerably less than when His231 was mutated (*21*). Though these residues in the secondary coordination sphere are not exactly in the same position in space with respect to zinc, they provide the same kind of interactions and accordingly suggest that the catalytic mechanism of all three will be similar.

Comparison of Tyr350Ala Molecule A with the Thermolysin−Inhibitor Complex. The superposition of molecule A of Tyr350Ala on the structure with the thermolysin−inhibitor complex model shows interesting similarity (Figure 5c). The sulfate ion superimposes on the phosphate ion of the inhibitor very well and has similar interactions. The sulfate ion mimics the transitional tetrahedral geometry of the scissile peptide bond carbonyl carbon as in the case of the BoNT/B structure with a sulfate ion (*29*). The sulfur atom of sulfate imitates the carbonyl carbon of the scissile bond. O1 and O3 of the sulfate ion at distances 2.40 and 2.61 Å from zinc respectively mimic the positions of the scissile peptide bond carbonyl oxygen and displaced nucleophilic water. The O3 of sulfate makes hydrogen bonds with Glu212 OE1 and OE2 at 2.80 and 2.53 Å, respectively, and the O1 of sulfate with Glu250 OE1 at a distance of 2.84 Å. O4 of the sulfate represents the amide nitrogen of the scissile bond which is in hydrogen-bonding distance to Glu159 O as in thermolysin (Figure 6). This comparison also helps in mapping the S1′ subsite where P1′ (Ile181) of the substrate will bind. The side chain of P1′ occupies a cavity lined by Thr159, Phe191, Gln203, Leu207, Thr208, and Tyr356. The molecular surface in this region is shown in Figure 7.

From the models of thermolysin and neprilysin with their inhibitors it is proposed that Tyr350 and Arg347 interact with the carbonyl oxygens of P1 and P1′ of the substrate (here SNAP-25). When these interactions are not available, the catalytic activity is affected. This model could not be directly compared with the BoNT/A-SNAP-25 complex, since the scissile bond carbonyl oxygen is >6.5 Å away from zinc, maybe because of the absence of Tyr required for stabilization of the substrate (*7*). The overall comparison suggests that the substrate docking at the active site as well as the orientation of the peptide bond may be similar in all botulinum neurotoxins and tetanus neurotoxin even though the substrate and scissile bonds are different.

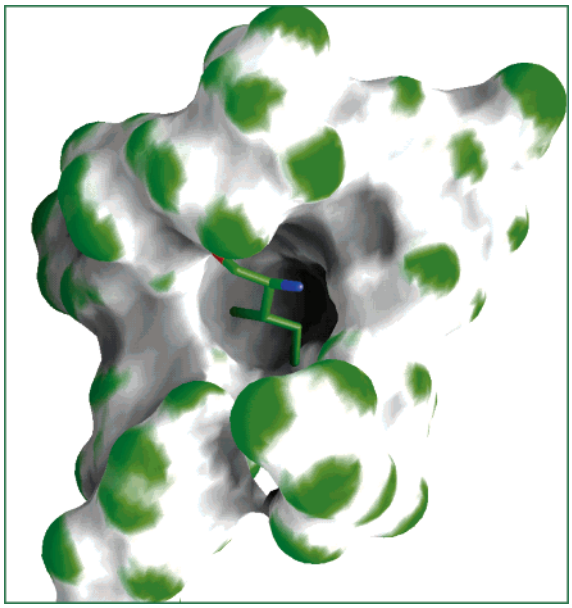

FIGURE 7: The molecular curvature surface formed by residues lying within 5 Å radius from P1′ (Ile181 of the substrate) is shown. The cavity occupied by the side chain of P1′ is lined by Thr159, Phe191, Gln203, Thr208, and Tyr356. Ile181 of the substrate alone is shown as a rod model. This figure was generated with GRASP (*35*).

CONCLUSION

Overall, this study identifies the crucial role of nucleophilic water's presence, position, and polarizing ability. Since its position is determined by the conserved residues in the active site, an effective inhibitor common to all BoNTs can be

designed to block these conserved residues in combination. In this report we have also shown that subtle changes in the electrostatic potential due to changes on surface charges distant from the active site could drastically change the environment and the activity of the enzyme without altering the conformation. Glu335Gln is a persistent apoenzyme and could be a candidate for a recombinant vaccine. The double mutant Glu212Gln/Glu335Gln or Glu212Gln/Tyr350Ala could be a better target since both will be completely inactive with the former devoid of zinc. The residual activity on Glu335Gln also points to the fact that the nucleophilic water could still be attacking the scissile peptide bond carbonyl carbon in a way that is yet to be understood. In summary, we have shown that Arg347 and Tyr350 play a crucial role in substrate stabilization while Glu249 plays a supportive role in catalysis. Glu335, though in the secondary coordination sphere, plays a dominant role in changing the environment of the active site and the catalytic activity.

ACKNOWLEDGMENT

We thank Drs. A. Saxena and M. Becker for providing beam time at the National Synchrotron Light Source, Brookhaven National Laboratory, J. Romeo for technical assistance, and Drs. S. Eswaramoorthy and D. Kumaran for helpful discussions.

REFERENCES

1. Sollner, T., Whiteheart, S. W., Brunner, M., Erdjument-Bromage, H., Geromanos, S., Tempst, P., and Rothman, J. E. (1993) SNAP receptors implicated in vesicle targeting and fusion, *Nature 362*, 318−324.
2. Schiavo, G., Matteoli, M., and Montecucco, C. (2000) Neurotoxins affecting neuroexocytosis, *Physiol. Rev. 80*, 717−766.
3. Agarwal, R., Eswaramoorthy, S., Kumaran, D., Binz, T., and Swaminathan, S. (2004) Structural analysis of botulinum neurotoxin type E catalytic domain and its mutant Glu212→Gln reveals the pivotal role of the Glu212 carboxylate in the catalytic pathway, *Biochemistry 43*, 6637−6644.
4. Hanson, M. A., and Stevens, R. C. (2000) Cocrystal structure of synaptobrevin-ll bound to botulinum neurotoxin type B at 2.0 Å resolution, *Nat. Struct. Biol. 7*, 687−692.
5. Rao, K. N., Kumaran, D., Binz, T., and Swaminathan, S. (2005) Structural studies on the catalytic domain of clostridial tetanus toxin, *Toxicon* (in press).
6. Segelke, B., Knapp, M., Kadkhodayan, S., Balhorn, R., and Rupp, B. (2004) Crystal structure of *Clostridium botulinum* neurotoxin protease in a product-bound state: Evidence for noncanonical zinc protease activity, *Proc. Natl. Acad. Sci. U.S.A. 101*, 6888−6893.
7. Breidenbach, M. A., and Brunger, A. (2004) Substrate recognition strategy for botulinum neurotoxin serotype A, *Nature 432*, 925−929.
8. Lebeda, F. J., and Olson, M. A. (1995) Structural predictions of the channel-forming region of botulinum neurotoxin heavy chain, *Toxicon 33*, 559−567.
9. Kurazono, H., Mochida, S., Binz, T., Eisel, U., Quanz, M., Grebenstein, O., Wernars, K., Poulain, B., Tauc, L., and Niemann, H. (1992) Minimal essential domains specifying toxicity of the light chains of tetanus toxin and botulinum neurotoxin type A, *J. Biol. Chem. 267*, 14721−14729.
10. Rossetto, O., Schiavo, G., Montecucco, C., Poulain, B., Deloye, F., Lozzi, L., and Shone, C. C. (1994) SNARE motif and neurotoxins, *Nature 372*, 415−416.
11. Vaidyanathan, V. V., Yoshino, K., Jahnz, M., Dorries, C., Bade, S., Nauenburg, S., Niemann, H., and Binz, T. (1999) Proteolysis of SNAP-25 isoforms by botulinum neurotoxin types A, C, and E: Domains and amino acid residues controlling the formation of enzyme−substrate complexes and cleavage, *J. Neurochem. 72*, 327−337.
12. Washbourne, P., Pellizzari, R., Baldini, G., Wilson, M. C., and Montecucco, C. (1997) Botulinum neurotoxin A and E require the SNARE motif in SNAP-25 for proteolysis, *FEBS Lett. 418*, 1−5.
13. Binz, T., Bade, S., Rummel, A., Kollewe, A., and Alves, J. (2002) Arg^{362} and Tyr^{365} of the botulinum neurotoxin type A light chain are involved in transition state stabilization, *Biochemistry 41*, 1717−1723.
14. Rigoni, M., Caccin, P., Johnson, E. A., Montecucco, C., and Rossetto, O. (2001) Site-directed mutagenesis identifies active-site residues of the light chain of botulinum neurotoxin type A, *Biochem. Biophys. Res. Commun. 288*, 1231−1237.
15. Rossetto, O., Caccin, P., Rigoni, M., Tonello, F., Bortoletto, N., Stevens, R. C., and Montecucco, C. (2001) Active-site mutagenesis of tetanus neurotoxin implicates TYR-375 and GLU-271 in metalloproteolytic activity, *Toxicon 39*, 115−1159.
16. Yamasaki, S., Hu, Y., Binz, T., Kalkuhl, A., Kurazono, H., Tamura, T., Jahn, R., Kandel, E., and Niemann, H. (1994) Synaptobrevin/vesicle-associated membrane protein (VAMP) of *Aplysia californica*: structure and proteolysis by tetanus toxin and botulinal neurotoxins type D and F, *Proc. Natl. Acad. Sci. U.S.A. 91*, 4688−4692.
17. Li, L., Binz, T., Niemann, H., and Singh, B. R. (2000) Probing the mechanistic role of glutamate residues in the zinc-binding motif of type A botulinum neurotoxin light chain, *Biochemistry 39*, 2399−2405.
18. Matthews, B. W. (1988) Structural basis of the action of thermolysin and related zinc peptides, *Acc. Chem. Res. 21*, 333−340.
19. Marie-Claire, C., Ruffet, E., Antonczak, S., Beaumont, A., O'Donohue, M., Roques, B. P., and Fournie-Zaluski, M. C. (1997) Evidence by site-directed mutagenesis that arginine 203 of thermolysin and arginine 717 of neprilysin (neutral endopeptidase) play equivalent critical roles in substrate hydrolysis and inhibitor binding, *Biochemistry 36*, 13938−13945.
20. Marie-Claire, C., Ruffet, E., Tiraboschi, G., and Fournie-Zaluski, M. C. (1998) Differences in transition state stabilization between thermolysin (EC 3.4.24.27) and neprilysin (EC 3.4.24.11), *FEBS Lett. 438*, 215−219.
21. Beaumont, A., O'Donohue, M. J., Paredes, N., Rousselet, N., Assicot, M., Bohuon, C., Fournie-Zaluski, M. C., and Roques, B. P. (1995) The role of histidine 231 in thermolysin-like enzymes. A site-directed mutagenesis study, *J. Biol. Chem. 270*, 16803−16808.
22. de Kreij, A., van den Burg, B., Venema, G., Vriend, G., Eijsink, V. G., and Nielsen, J. E. (2002) The effects of modifying the surface charge on the catalytic activity of a thermolysin-like protease, *J. Biol. Chem. 277*, 15432−15438.
23. Agarwal, R., Eswaramoorthy, S., Kumaran, D., Dunn, J. J., and Swaminathan, S. (2004) Cloning, high level expression, purification and crystallization of the full length *Clostridium botulinum* neurotoxin type E light chain, *Protein Expression Purif. 34*, 95−102.
24. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
25. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszwewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Somonsom, T., and Warren, G. L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
26. Jones, T. A., Zou, J., Cowtan, S., and Kjeldgaard, M. (1991) Improved methods in building protein models in electron density map and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.
27. Lipscomb, W. N., and Strater, N. (1996) Recent advances in zinc enzymology, *Chem. Rev. 96*, 2376−2433.
28. Swaminathan, S., and Eswaramoorthy, S. (2000) Structural analysis of the catalytic and binding sites of *Clostridium botulinum* neurotoxin B, *Nat. Struct. Biol. 7*, 693−699.
29. Swaminathan, S., Eswaramoorthy, S., and Kumaran, D. (2004) Structure and enzymatic activity of botulinum neurotoxins, *Movement Disorders 19* (Suppl. 8), S17−S22.

30. Eswaramoorthy, S., Kumaran, D., Keller, J., and Swaminathan, S. (2004) Role of metals in the biological activity of *Clostridium botulinum* neurotoxins, *Biochemistry 43*, 2209−2216.
31. Mock, W. L., and Aksamawati, M. (1994) Binding to thermolysin of phenolate-containing inhibitors necessitates a revised mechanism of catalysis, *Biochem. J. 302*, 57−68.
32. Lister, S. A., Wetmore, D. R., Roche, R. S., and Codding, P. W. (1996) E144S Active-site mutant of the *Bacillus cereus* thermolysin-like neutral protease at 2.8 Å Resolution, *Acta Crystallogr. D52*, 543−550.
33. Li, L., and Singh, B. R. (2000) Role of zinc binding in type A botulinum neurotoxin light chain's toxic structure, *Biochemistry 39*, 10581−10586.
34. Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of proteins, *J. Appl. Crystallogr. 24*, 946−950.
35. Nicholls, A., Sharp, K., and Honig, B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins 11*, 281−296.

BI050253A

Appendix 3

Structural Analysis of Botulinum Neurotoxin Serotype F Light Chain: Implications on Substrate Binding and Inhibitor Design†

Rakhi Agarwal,‡ Thomas Binz,§ and Subramanyam Swaminathan*,‡

Biology Department, Brookhaven National Laboratory, Upton, New York 11973, and Department of Biochemistry, Medizinische Hochschule Hannover, 30623 Hannover, Germany

Received May 27, 2005; Revised Manuscript Received July 11, 2005

ABSTRACT: The seven serologically distinct *Clostridium botulinum* neurotoxins (BoNTs A−G) are zinc endopeptidases which block the neurotransmitter release by cleaving one of the three proteins of the soluble *N*-ethylmaleimide-sensitive-factor attachment protein receptor complex (SNARE complex) essential for the fusion of vesicles containing neurotransmitters with target membranes. These metallopeptidases exhibit unique specificity for the substrates and peptide bonds they cleave. Development of countermeasures and therapeutics for BoNTs is a priority because of their extreme toxicity and potential misuse as biowarfare agents. Though they share sequence homology and structural similarity, the structural information on each one of them is required to understand the mechanism of action of all of them because of their specificity. Unraveling the mechanism will help in the ultimate goal of developing inhibitors as antibotulinum drugs for the toxins. Here, we report the high-resolution structure of active BoNT/F catalytic domain in two crystal forms. The structure was exploited for modeling the substrate binding and identifying the S1′ subsite and the putative exosites which are different from BoNT/A or BoNT/B. The orientation of docking of the substrate at the active site is consistent with the experimental BoNT/A-LC:SNAP-25 peptide model and our proposed model for BoNT/E-LC:SNAP-25.

Clostridium botulinum produces seven antigenically distinct neurotoxins (BoNT/A−G)[1] recognized as the most potent biological toxins (*1*). BoNTs cause the disease botulism, a neuromuscular disorder characterized by flaccid paralysis that is due to the blockage of neurotransmitter release (*2*). Botulism is also caused by other members of the genus such as *Clostridium baratii*. There are a few case reports on botulism due to BoNT/F secreted by the *C. baratii* in adults and infants (*3*−*5*). BoNTs are released as inactive molecules of 150 kDa and are cleaved by exogenous or endogenous proteases into active dichain molecules with an N-terminal light chain (LC, 50 kDa) and a C-terminal heavy chain (HC, 100 kDa) held together by a disulfide bond (*6*). The C-terminal domain of the HC mediates binding of the toxin to the specific neuronal receptors, and the N-terminal domain enables the catalytically active LC to translocate to the cytosol, where it recognizes and cleaves one of three SNARE proteins. Though LCs of BoNTs share significant sequence homology (30−60%) and structural similarity, they selectively cleave specific SNARE proteins. BoNT/A, -C, and -E cleave SNAP-25, and BoNT/B, -D, -F, and -G and tetanus toxin (TeNT) cleave VAMP (also known as synaptobrevin) (*7*). BoNT/C is unique since it also cleaves syntaxin. VAMP is a tail anchored membrane protein of exocytotic vesicles. VAMP comprises a family of different isoforms, and neurotoxin sensitive isoforms are present in many non-neuronal tissues (*8*). However, clostridial neurotoxins do not act on non-neuronal cells due to the absence of cell surface neurotoxin receptors to internalize them (*7*, *9*). The LCs are unique zinc endopeptidases as to the requirement of unusually large segments of their substrates for optimal activity. This is reflected in the presence of additional recognition and binding sites, called exosites, on BoNTs besides the active site. All three SNARE proteins contain conserved nine residue segments, termed SSR. There are four (S1−S4) such SSRs in SNAP-25 and two each in VAMP (V1 and V2) and Syntaxin (X1 and X2) (*10*). At least one of the SSRs is required for the substrate cleavage in addition to the cleavage site (*11*, *12*). While the cleavage site is at the C-terminal side of SSRs in most of them, BoNT/F and -D are unique in that they cleave peptide bonds located between V1 and V2. BoNT/F cleaves Gln58−Lys59 while BoNT/D cleaves the adjacent peptide bond, Lys59−Leu60. Since mutation of both V1 and V2 residues affects BoNT/F activity, it may require both V1 and V2 for activity unlike BoNT/D which needs only V1 (*12*).

† Research supported by the U.S. Army Medical Research Acquisition Activity (Award No. DAMD17-02-2-0011) under DOE Prime Contract No. DE-AC02-98CH10886 with Brookhaven National Laboratory. T.B. was supported by Grant RGY0027/2001 from Human Frontier Science Program.

* Author to whom correspondence should be addressed. E-mail: swami@bnl.gov. Tel: 1-(631)-344-3187. Fax: 1-(631)-344-3407.

‡ Brookhaven National Laboratory.

§ Medizinische Hochschule Hannover.

[1] Abbreviations: BoNT, botulinum neurotoxin; TeNT, tetanus neurotoxin; SNARE, soluble *N*-ethylmaleimide-sensitive-factor attachment protein receptor; SSR, SNARE Secondary Recognition; LC, light chain; SNAP-25, 25 kDa synaptosome-associated protein; VAMP, vesicle-associated membrane protein; USAMRIID, United States Army Medical Research Institute of Infectious Diseases; PCR, polymerase chain reaction; GST, glutathione sulfotransferase; PMSF, phenylmethylsulfonyl fluoride; HEPES, *N*-[2-hydroxyethyl]piperazine-*N*′-[2-ethanesulfonic acid]; rmsd, root-mean-square deviation.

10.1021/bi0510072 CCC: $30.25
Published on Web 08/10/2005

Due to the unique substrate requirement and stringent cleavage selectivity, the development of a multivalent inhibitor against all seven serotypes seems unlikely. Thus, the current need is to understand the catalytic mechanism of each one of them, to define the variable substrate specificity and selective cleavages by them. Knowledge of enzyme exosites along with the residues of substrate involved in enzyme−substrate complex formation is a prerequisite for drug design for botulism. The docking of the substrate at the active site needs to be studied more precisely to understand the molecular mechanism of catalysis brought about by these toxins in order to develop inhibitors. To achieve this goal, high-resolution crystal structures of catalytic domains of all serotypes and their complexes with the substrates or substrate analogues would be a priority. So far only two complex structures are available (*13*, *14*).

Here we report the expression, purification, and crystal structure analysis of BoNT/F-LC in two crystal forms. This is the first structural report for any functional fragment of BoNT/F. We also compare BoNT/F-LC with the available crystal structures of BoNT/A, -B, and -E and TeNT. We have identified the S1′ subsite and the putative exosites in the toxin. The suggested orientation of docked substrate at the active site is consistent with thermolysin-inhibitor, BoNT/A-LC:SNAP-25 peptide, and our proposed model for the BoNT/E-LC:SNAP-25 complex structures (*13*, *15*, *16*).

MATERIALS AND METHODS

Cloning, Expression, and Purification of BoNT/F-LC. The pFB4 vector possessing the full length of BoNT/F-LC (Met1−Lys439; identical to strain NCTC 10821; GenBank accession number X81714) with 6xHis tag was used as template. This has four extra residues at the C-terminus. The LC encoding segment was PCR amplified using the forward primer (5′ ATG ACC ATG GGA ATG CCA GTT GTA A 3′) bearing a *Nco*I restriction site at the 5′ end and the reverse primer (5′ GAT GCT CGA GCC CGG GAG TTG GCG G 3′) bearing a *Xho*I site at the 3′ end. The PCR products were digested with the *Nco*I and *Xho*I restriction enzymes and ligated to the pET-28b vector between similar sites. The nucleotide sequence was confirmed for the entire BoNT/F-LC gene in both strands by the Big Dye termination Cycle Sequencing (Applied Biosystems). The plasmid pET-28b-LC was transformed to BL21 (DE3) cells for expression of the protein.

The expression and purification methods are similar to those for BoNT/E-LC (*17*). The 2xYT medium (1.6% bactotryptone, 1.0% bacto-yeast extract, and 0.5% NaCl) containing 100 μg/mL kanamycin was inoculated with the freezer stock of BL21(DE3) cells containing the pET-F-LC vector. Cells were grown in a shaking incubator at 37 °C until A_{600} reached 0.6−0.8. At this point, 1 mM IPTG was added and the cell growth was continued for an additional 14−16 h at 20 °C. Cells were then harvested for the protein preparation, and 30 mL of lysis buffer containing 50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 5 mM benzamidine, 0.5 mM PMSF, and 1 μg/mL pepstatin A was added supplemented with two tablets of protease inhibitor cocktail (Roche) and 0.5 mg/mL lysozyme (Sigma), 2 mL of Bugbuster (Novagen), and 6 mM iodoacetamide. The suspension was incubated at room temperature for 20−30 min, and then 2 μL of benzonase was added and the suspension was incubated for an additional 10 min. The suspension was centrifuged at 12 000 rpm for 30 min and the supernatant saved. The supernatant was allowed to mix with Ni-NTA agarose (Qiagen) prewashed with the phosphate buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl) for 1 h, poured into a column, and washed with 100 mL of phosphate buffer. The elutions were done with increasing concentrations of imidazole (10−250 mM) in phosphate buffer. The protein fractions eluted in the 250 mM imidazole concentration were pooled together and purified further in a size exclusion column (Superdex-75). The protein was concentrated to 10 mg/mL using Centricon for crystallization purpose. The protein concentration was measured by absorbance at 278 nm using a Perkin-Elmer spectrophotometer.

Expression and Purification of GST-Tagged VAMP. The pGEX-2T vector encoding VAMP2 (aa 1−115; human) with an N-terminal GST tag was kindly provided by J. Schmidt, USAMRIID. The GST-VAMP fusion protein was expressed in BL21 cells and purified using GST-resin as per user's manual (Novagen).

Enzymatic Activity of BoNT/F-LC. The enzymatic activity of BoNT/F-LC was assayed on its substrate VAMP. The assay was performed at 37 °C in 20 μL total volume (20 mM Hepes buffer, pH 7.4, 2 mM DTT, and 10 μM Zn acetate) containing 10 nM enzyme and 5 μM GST-VAMP. The hydrolytic curve (data not shown) for the light chain (2 nM) was determined by analyzing the cleavage of 5 μM substrate (GST-VAMP) at various time intervals (0, 1, 2, 3, 5, 10, 15, 30, 45, 60, 90, 120, and 180 min). The reaction was stopped by adding the 3x SDS−PAGE sample buffer containing 1 mM EDTA. The evaluation of cleavage was based on the appearance/disappearance and intensity of uncleaved substrate and product in 4−20% gradient Tris-glycine SDS−PAGE gels. Quantification of the digestion was performed by densitometry analysis of the cleaved and uncleaved fragments in Coomassie stained gel.

Screening of Mercurial Compounds as Inhibitors. The inhibitor assay was performed as above with two concentrations (500 and 1000 μM) of five mercurial compounds individually added to the assay buffer and incubated for 40 min at room temperature before adding the substrate. The reaction was stopped after 30 min by adding the 3x SDS−PAGE sample buffer containing 1 mM EDTA. The efficiency of cleavage was calculated as percentage of cleavage by the untreated enzyme as detailed above. The most promising compound was assayed again at finer grid points.

Crystallization and Data Collection. The initial crystallization screening was carried out by the sitting drop vapor diffusion method using the high-throughput crystal screen (Hampton Research) and TECAN crystallization robot. Microcrystals were obtained using a precipitant containing 0.1 M $CdCl_2$, 0.1 M Na-acetate pH 4.6, and 30% PEG 400 at room temperature in 2−3 days. Diffraction quality crystals were obtained by microseeding and varying the concentration of $CdCl_2$. Two morphologically distinct crystal forms, hexagonal (form I) and rod (form II) shaped, appeared under the same crystallization conditions. Preliminary analysis showed that both forms are in the space group *C*2. Form I crystals contain two molecules per asymmetric unit while form II crystals have one molecule per asymmetric unit. A

Table 1: Crystal Data and Refinement Statistics

form	Crystal Data Details
I	$a = 173.40$ Å, $b = 53.24$ Å, $c = 113.87$ Å, $\beta = 119.2°$; space group $C2$
II	$a = 63.41$ Å, $b = 79.46$ Å, $c = 89.66$ Å, $\beta = 110.0°$; space group $C2$

	form I	form II
wavelength (Å)	1.1	1.1
resolution (Å)	50−1.8	50−2.0
no. of reflections	69805	26560
I/σ	7.3	11.1
R-merge (outer shell)	0.09 (0.42)	0.08 (0.21)
redundancy	3.0	6.5
completeness (%)	82.9 (47.4)	93.7 (62.7)
Refinement Statistics		
resolution used (Å)	50−1.8	50−2.0
molecules/asymmetric unit	2	1
no. of reflections	64843	25768
completeness (%)	76.7	90.9
R-factor	0.24	0.23
R-free	0.27	0.28
no. of protein atoms	6338	3264
no. of heteroatoms	9	3
no. of water molecules	241	167
Ramachandran plot most favored region (additionally allowed) (%)	88.4 (10.4)	90 (9.2)
PDB id code	2A97	2A8A

Table 2: Interatomic Distances between Various Conserved Residues

	distance (Å)		distance (Å)
Zn−H227 NE2	2.00	Nu water−R365 NH2	7.58
Zn−H231 NE2	2.22	Nu water−H231 NE2	3.21
Zn−E266 OE1	2.20	E353 OE1−H227 ND1	3.44
Zn−E250 OE2	2.96	E353 OE2−R365 NH1	4.01
Zn−Nu water	2.13	E265 OE1−H231 ND1	2.82
Nu water−E228 OE1	3.74	E265 OE1−H234 ND1	2.63
Nu water−E228 OE2	4.42	Y368-OH−Zn	4.37
Nu water−Y368-OH	4.64	Y368-OH−E266 OE1	2.87
Nu water−E266 OE1	3.49	Y368-OH−E266 OE2	3.71
Nu water−E266 OE2	3.02	R365 NH2−Zn	6.62

self-rotation function calculation showed the presence of a noncrystallographic 2-fold symmetry in form I crystals (*18*).

Crystals were directly mounted on a nylon loop and flash frozen immediately by plunging into liquid nitrogen. X-ray diffraction data were collected at the SGX-CAT beam line of the APS for form I crystals and at the X29 beamline of the NSLS for form II crystals. All data sets were processed with DENZO and scaled and merged using SCALPACK (*19*). The data collection statistics are given in Table 1.

Structure Determination. Crystal structures of forms I and II were determined using the molecular replacement method with BoNT/B-LC (1F82) as search model (*18*). ARP/wARP was used to build the model (*20*). About 80% of the model was built automatically and the rest manually using "O" (*21*). Refinement was carried out using CNS (*22*). Fifteen C-terminal residues could not be located in the electron density maps in both crystal forms and were presumed to be disordered. In addition, form I crystals lacked electron density for the loop regions Asp205−Thr214 and Glu249−Ile262 in both monomers while form II crystals lacked only for Asp205−Thr214. These loops have a tendency to be disordered in all LCs (*14*, *23*, *24*). Refinement details are given in Table 1.

Modeling of Substrate on BoNT/F-LC. Two crystal structures of botulinum neurotoxin catalytic domain−substrate complex are available (*13*, *14*). The two LCs have different substrates, and the mode of binding is also different. To model BoNT/F-LC with its substrate, both models were tried. First BoNT/B-LC with VAMP peptide was aligned with BoNT/F-LC by least-squares fit. In this process the disordered loops were not included. Similar alignment was done with BoNT/A (mutant):SNAP-25 peptide. In the latter case, the SNAP-25 peptide was converted into VAMP peptide via mutation in "O".

RESULTS AND DISCUSSION

Cloning, Expression, Purification, and Catalytic Activity of BoNT/F-LC. The BoNT/F catalytic domain gene from pFB4 plasmid has been recloned into a pET-28b vector with C-terminal 6xHis tag (pET-F-LC). Using pET-F-LC BoNT/F-LC has been overexpressed yielding >25 mg/L cell culture. The solubility of the protein was extremely good, and we could concentrate it to >10 mg/mL in HEPES buffer. The protein is stable when stored in the same buffer between −80 and −20 °C.

VAMP2 (aa 1−115) with the N-terminal GST tag was used for assaying enzymatic activity. The GST-tag (25 kDa) does not interfere with the catalytic activity (*12*). Incubation of GST-VAMP (∼38 kDa) with the BoNT/F-LC led to cleavage of the Gln58−Lys59 peptide bond and produced ∼32 and ∼6 kDa fragments. LC, 10 nM, was able to completely cleave 5 μM VAMP in 30 min of incubation at 37 °C (data not shown).

Description of BoNT/F-LC Structures. BoNT/F-LC has the fold typical of the available botulinum toxin catalytic domains. The protein has one zinc ion per molecule. Form I and form II crystal structures are very similar. The rmsd between the two monomers in form I is 0.40 Å, and that between form II molecule and either one of form I monomers is 0.43 Å. The loop region Asp205−Thr214 is disordered and not modeled in both structures. Loop Glu249−Ile262 disordered and hence not modeled in form I is ordered in the form II crystal. The C-terminal residues (Val422−Lys439) are not modeled due to poor electron density in both structures. The absence of the C-terminal residues and the disordered region (Glu249−Ile262) is not due to autolysis as shown by SDS−PAGE analysis on crystals of both forms (data not shown) (*24*, *25*). In form I crystals the two molecules (A and B) in the asymmetric unit are associated via a noncrystallographic 2-fold symmetry. One β-strand from each molecule is held together by a cadmium ion, which is interacting with OD1 and OD2 of Asp383 of both monomers with distances of 2.27, 2.28, 2.72, and 2.52 Å. Other interactions are primarily through loop regions. Residues 7−13 of molecule A interact with residues 394−396 of molecule B and vice versa. The role of cadmium ions is probably to stabilize the packing and thus help in crystallization. Though the molecule exists as a dimer in the crystal structure, the buried surface area of 1680 Å^2 (about 5% of the total area) on dimerization indicates that it is not a biological dimer. This notion is supported by results of size exclusion chromatography experiments where the peak

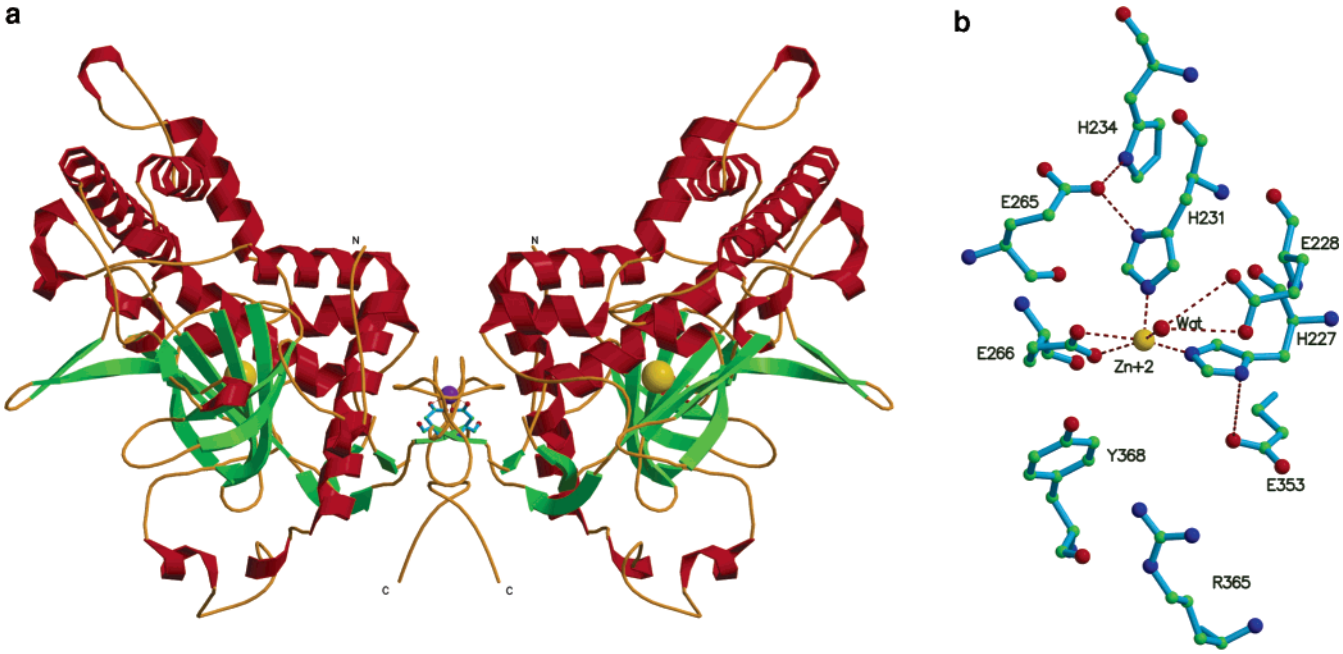

FIGURE 1: (a) Ribbon representation of the dimer formed via crystallographic 2-fold in form II crystals. The dimerization is similar in form I, but via a noncrystallographic 2-fold. The dimer formation is aided by a Cd ion (shown in magenta) coordinated by the carboxylate side chain of Asp383, shown in ball-and-stick model. The active-site zinc is shown in yellow. The dimerization may be an artifact of crystal packing. This figure was produced with MOLSCRIPT (*41*). (b) The active site of BoNT/F-LC. The zinc ion is coordinated by His227, His231, Glu266, and a nucleophilic water (wat) which is hydrogen bonded to Glu228. The interatomic distances are given in Table 2.

due to the protein appeared at 52 kDa when eluted at a concentration of 1−2 mg/mL. In both the molecules the active sites are facing outside the dimeric interface and exposed to solvent (Figure 1a). The dimeric association is different in all the clostridial neurotoxin light chain structures reported so far (*23*, *24*, *26*). Whether or not the dimerization has any biological significance is not yet clear. In form II crystals similar dimerization is achieved via the crystallographic 2-fold, also with a cadmium ion bridging the molecules. Since all monomers and their interatomic distances in both crystal forms agree within experimental errors, the discussion hereafter pertains to form II crystals as the model is more complete (Table 1).

The BoNT/F-LC structure is similar to BoNT/A, -B, and -E and TeNT LCs with respect to the core conformation involving the helices and strands though large variations exist in their loop conformations. Of the full-length catalytic domain structures available so far (BoNT/F and -E and TeNT), the C-terminal region is visible only in BoNT/E-LC (*23*, *26*). In BoNT/F-LC, the active site is well formed and has interactions typical of zinc-binding motifs as in other clostridial neurotoxins. The active site zinc is coordinated by His227 NE2, His231 NE2, Glu266 OE1, and a nucleophilic water molecule (Figure 1b). The latter is kept in its position via Glu228.

Counterparts of the remaining depicted active site residues are present in analogous arrangement in all clostridial neurotoxins crystallized to date. This argues for the functional similarity of these residues. The function of several active sites residues has been studied by mutational analysis (*16*, *27*−*30*). Thus, Glu265 appears to be important for maintaining the orientation and protonation state of His231. Glu353 likely fulfills the same function on His227. Arg365 probably helps in stabilization of the P1′ carbonyl oxygen of the scissile bond of the substrate to give proper orientation. Tyr368 may play a major role in transition state stabilization as mutation of the equivalent residue in BoNT/E (Tyr350) to alanine inactivated the enzyme (*16*). However, moderate differences in phenotypes have been observed for mutations of analogous residues in different clostridial neurotoxins.

Putative Substrate Binding Sites for BoNT/F-LC. Two crystal structures of BoNT LCs with their substrates are available (*13*, *14*). The two toxins have different substrates and could hence represent the mode of binding of the particular substrate to the catalytic domain. In both models, part of the substrate runs close to the β strand formed by residues in the region 160−170 (BoNT/A). In all LC structures and the holotoxins, this region forms a β strand. In our modeling the transition state by comparing BoNT/B with a sulfate ion with thermolysin, we had pointed out that residues in this strand might interact with the substrate (*31*). A similar conclusion has been drawn also from mutational studies in BoNT/E-LC (*16*). However, while the substrate runs antiparallel to this strand to form a β sheet in the BoNT/A:SNAP-25 complex, in the BoNT/B:VAMP complex, it runs close to the β strand but is in the same direction as the strand. This difference could be because of the difference in substrate type and/or the serotype. This raises an interesting possibility that all LCs which have VAMP as substrate will bind similarly to BoNT/B and those with SNAP-25 as substrate similarly to BoNT/A. But a reversal of direction is also a possibility for the same substrate when it binds to a different serotype. To model the substrate binding to BoNT/F-LC we took both possibilities into account. Accordingly, the VAMP peptide was modeled similarly to SNAP-25 in the BoNT/A:SNAP-25 peptide complex (Type 1). It was also modeled by tracing the chain in the reverse direction (Type 2). Both models were analyzed with respect to biochemical and mutational studies (*12*, *32*).

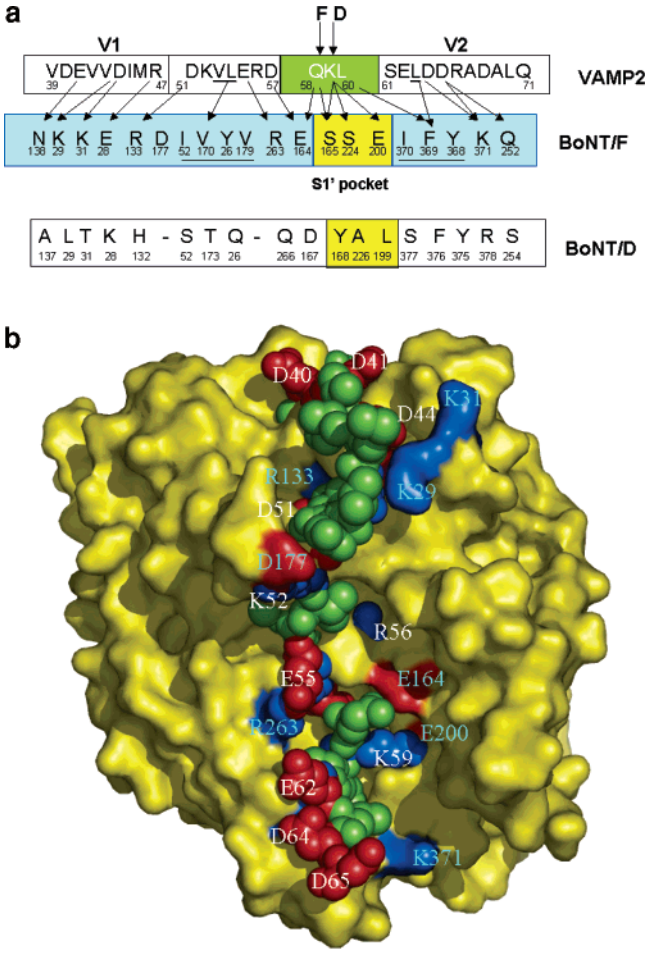

FIGURE 2: (a) Proposed VAMP−BoNT/F interactions. The scissile peptide bonds for BoNT/F and -D are shown with arrow marks. The residues involved in S1′ subsite are shown in yellow. The three residues forming the S1′ subsite will form hydrogen bonds with the P1′ residue, Lys59. The SSR motifs, V1 and V2, are also shown. The putative interactions between VAMP and BoNT/F are shown with arrow marks. Most of the interactions are between residues with complementary charges. Residues involved in the putative hydrophobic interactions are underlined. Strong interactions of V1 and V2 residues with the enzyme explain the need of both motifs for optimal catalytic activity of BoNT/F. BoNT/D residues corresponding to interacting residues of BoNT/F are also shown. Among the interacting residues only Tyr368 and Phe369 (BoNT/F) are conserved in BoNT/D and -F. The putative S1′ subsite for BoNT/D is highlighted in yellow. To allow for the hydrophobic interactions between P1′ and the S1′ subsite, VAMP will have to slide up by one residue. (b) BoNT/F-LC:VAMP complex model. BoNT/F-LC is shown as molecular surface in yellow with charged residues interacting with VAMP in red (negative) and blue (positive). VAMP (39−65) is shown in sphere model. Charged residues are shown in red and blue, while the rest are shown in green. The complementary charges between VAMP and BoNT/F-LC are evident. Charged residues are labeled in white for VAMP and in cyan for BoNT/F-LC. Figures 2b, 3, 4, and 5 were prepared with PyMOL (*42*).

The cleavage site in VAMP2 by BoNT/F-LC lies between the two SSR sites, V1 and V2 (Figure 2a) (*12*). The cleavage of VAMP2 is susceptible to mutation of residues in both SSRs, though more so for V1 residues than V2. When Asp40, Glu41, and Asp44 of V1 were mutated, the activity was almost completely lost. When Asp64, Asp65, and Asp68 of V2 were mutated, the cleavage decreased by 40−60% (*12*). In type 1 alignment, Asp44 has interactions with two positively charged residues (Lys29 and Lys31 of the LC) and could form a salt bridge to stabilize binding. According to the model, i.e., basing on a similar progression of the substrate along the binding channel as found in BoNT/A: SNAP-25, while Glu41 could make a potential hydrogen bond with the side chain of Asn138 of BoNT/F, Asp40 does not appear to form interactions with the LC. In addition, Arg47 of VAMP is near Glu28 of BoNT/F. Asp51, Lys52, Glu55, and Arg56 of VAMP interact with Arg133, Asp177, Arg263, and Glu164 of BoNT/F, respectively. Hydrophobic residues Val53 and Leu54 sit in a hydrophobic pocket formed by Ile52, Val170, Tyr26, and Val179 of BoNT/F. On the C-terminal side of the cleavage site, Asp64 and Asp65 of V2 interact with Lys371 and Gln252 of BoNT/F (Figure 2b). All these interactions indicate strong binding of the substrate to BoNT/F and largely support the results from mutagenesis studies. Together, these interactions could be considered the exosite interactions of BoNT/F. These LC residues are not conserved in BoNT/D (Figure 2a), and this could be the reason for different effects on the activity due to mutation of SSR residues even though the two cleave adjacent peptide bonds (*12*, *33*, *34*). On the other hand, in type 2 alignment, residues with the same kind of charge clash between the substrate and LC. Accordingly, we conclude that VAMP binds BoNT/F in a direction similar to SNAP-25 to BoNT/A. It remains to be shown whether the different direction of substrate binding in BoNT/B is unique or applies to other LCs as well. In discussing the need for the presence of V1 for TeNT activity, it was assumed that the VAMP will bind similarly to BoNT/B (*26*). However, the reverse substrate orientation can at present not be excluded. At any rate, the different specificity of TeNT and BoNTs is due to variation in the charge distribution on the surface of the LCs which allows different segments of the same substrate to bind presenting the specific cleavage site at the active site. Figure 3 provides a comparison of the charge distribution on BoNT/F-, BoNT/B-, and TeNT-LCs, all cleaving the same substrate.

S1′ Pocket of BoNT/F-LC. The requirements on P2, P1, P1′, and P2′ residues with respect to cleavage by clostridial neurotoxins have been studied extensively (*32*, *35*−*37*). While variation of P1 is irrelevant to cleavage (except for BoNT/C (*37*)), P1′ is crucial. The S1′ subsite has been identified in BoNT/E (*16*). The P1′ residue, Ile181, in BoNT/E occupies a hydrophobic site formed by Thr159, Phe191, Gln203, Leu207, Thr208, and Tyr356. A similar modeling with BoNT/F-LC and thermolysin proposes the S1′ subsite of BoNT/F. In this case, the P1′ residue, Lys59, interacts with Ser165, Glu200, and Ser224 (Figure 4). In BoNT/A these residues correspond to Phe162, Thr192, and Thr219 and have the potential to interact with P1′ (Arg198 of SNAP25) with some rearrangement of the side chain orientations. A proper rotamer position of Lys59 of VAMP will allow its side chain to interact with the side chains of these three residues. Changing Lys59 of VAMP to either Arg or Ala rendered VAMP cleavage resistant (*32*). Though Arg is positively charged as Lys, the side chain is too big to be accommodated in this subsite and is detrimental to the activity. As for Ala, the lack of interactions with the subsite residues will produce the same effect. In our substrate complex model, P1 is exposed to the solvent and has minimal interactions with the LC. This may be the case for other complexes and might explain why the type of amino acid at P1 is not crucial for cleavage. For BoNT/F, P2 and P2′ of the substrate are crucial for cleavage. P2 is Asp57, and its

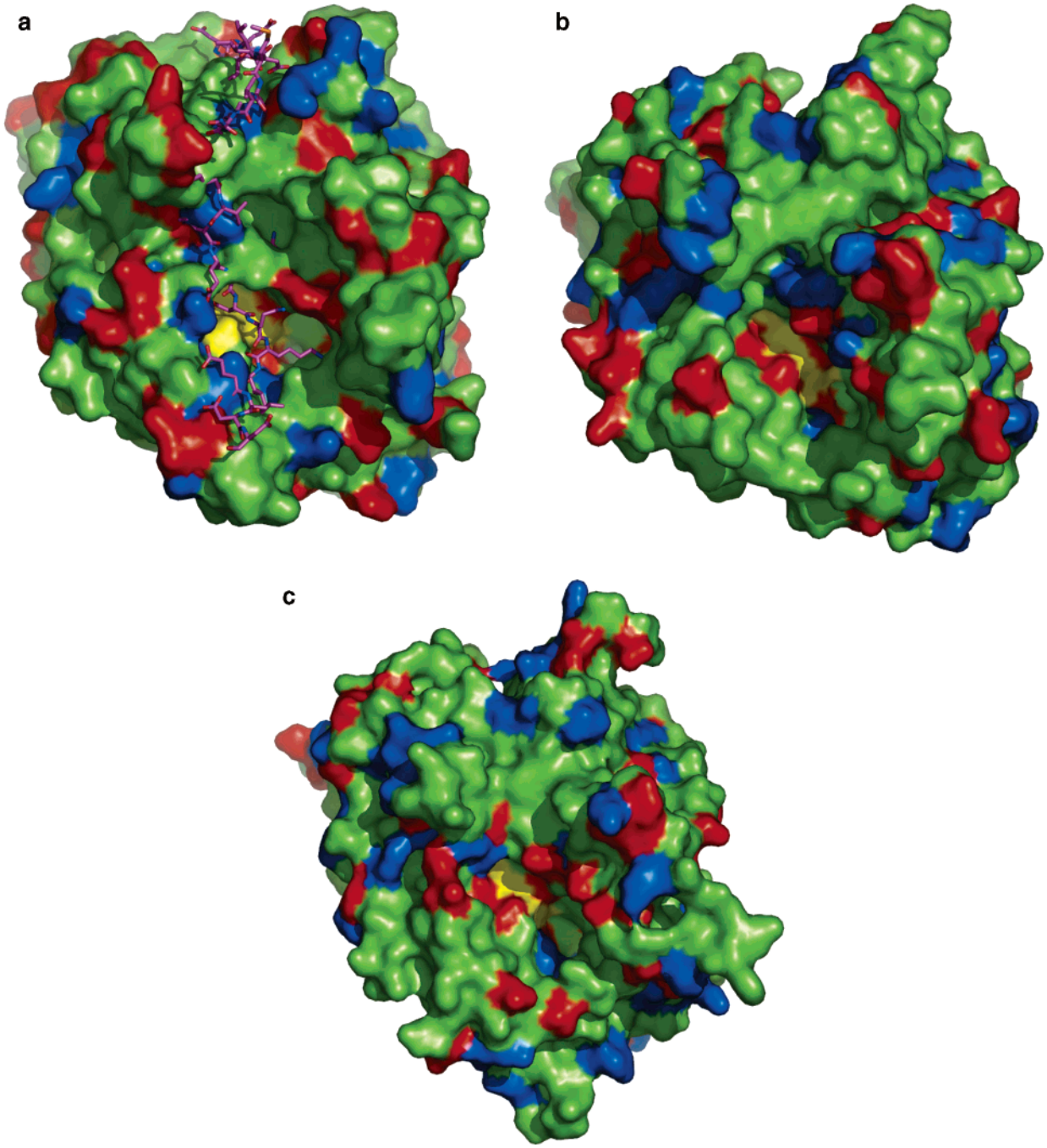

FIGURE 3: Comparison of the surface charges on (a) BoNT/F-LC, (b) TeNT-LC (PDB id 1YVG), and (c) BoNT/B-LC (PDB id 1F83). The three molecules are in the same orientation, and VAMP modeled on BoNT/F-LC is shown in stick model. Residues with positive and negative charges are shown in red and blue, respectively, and the active site region is shown in yellow. The variation in charge distribution explains the specificity for different cleavage sites. (Though BoNT/B and TeNT cleave the same peptide bond of VAMP, the optimal length of VAMP for activity is different. While TeNT requires both V1 and V2 SSR sites, BoNT/B requires only V2.)

side chain interacts with Arg263 of the LC probably via a salt bridge. Changing Asp57 to either Asn or Ala or Glu affected the activity. In the case of Glu the side chain is one bond length larger and might have steric clash. P2′ (Leu60) occupies a pocket formed by bulky hydrophobic residues Tyr368, Phe369, and Ile370. Change of the P2′ residue to Ala compromises the hydrophobic interactions and consequently affects the rate of substrate cleavage.

Mercury Compounds as Inhibitors of Botulinum Neurotoxins. The role of mercurial compounds on the activity of BoNTs has been investigated with respect to the interchain disulfide bridge connecting the heavy and light chains (*38*). However, the same compound had different effects on various serotypes. Ahmed et al. also have shown that similar mercurial compounds completely abolish the activity of BoNT/A-LC at 10−50 μM concentrations (*39*), and this is attributed to the modification of the thiol group of Cys164 which is in the vicinity of the active site. We chose to investigate the effect of mercury compounds on BoNT/F since there are two adjacent cysteines (Cys166 and Cys167) in the β strand that would form a β sheet with VAMP according to our enzyme−substrate complex model described in earlier sections. Cys167 of BoNT/F and Cys164 of BoNT/A are conserved. In addition, Cys364 is close to Arg365 and

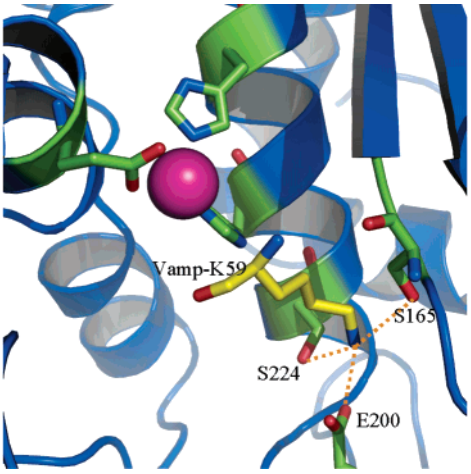

FIGURE 4: Close-up view of the active site and S1′ subsite. The active site zinc is shown in magenta, and the residues involved in the active and S1′ subsite are shown in stick model. The P1′ residue Lys59 is shown in yellow. The three potential hydrogen bonds are shown as dashed lines.

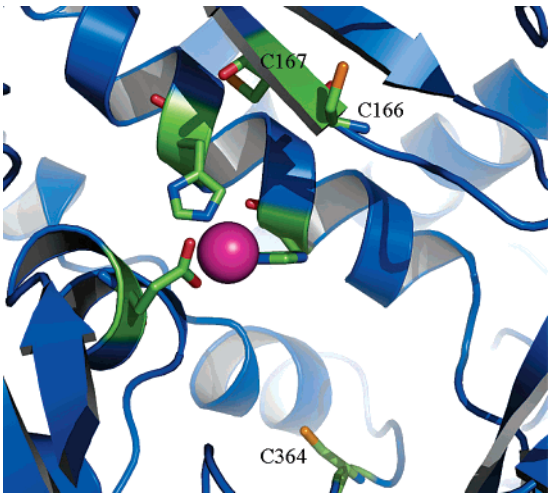

FIGURE 5: Cysteines near the active site. The three cysteines (166, 167, and 364) within 11 Å from the active site zinc are shown. Cys364 is close to Arg365 and Tyr368 which stabilize the transition state. The fourth cysteine (Cys429) was not modeled in the structure.

Tyr368 (Figure 5). The three mercury compounds we tried on BoNT/F-LC had varying effects in inhibiting the catalytic activity. As has been pointed out (*38*), this could be because of the modification of thiol groups or histidines near the active site or both. Future crystallographic investigations may help in understanding the role of mercury compounds in inhibiting the catalytic activity.

The Rationale for Inhibitor Development. BoNT/A, -E, and -C proteolytically cleave SNAP-25 while BoNT/B, -D, -F, and -G cleave VAMP. Small molecule inhibitors such as zinc chelators are relatively inefficient in blocking the activity of botulinum neurotoxins. This may be because either they do not bind tightly at the active site or they could be easily displaced by the substrate. In addition, small molecule zinc chelators could have undesirable side effects since they might inhibit other zinc proteins, with similar zinc binding motifs, essential for metabolism (*40*). In view of this, it would be advantageous to design inhibitors which would bind tightly covering a larger area of the enzyme. Mutations of residues in the light chain farther away from the active site either abolish or reduce the catalytic activity. The substrates are large peptides, and the enzyme−substrate models indicate strong binding of substrate segments in more than one site. This fact could be taken advantage of in developing peptidomimics. The competitive removal of the small substrate peptides by the larger substrate can be overcome by using inhibitors that will bind in more than one site. Another approach would be to block the S1 and S1′ subsites and exosites with two or three small molecules connected by linkers. Such molecules have been tried successfully in other cases before.

A further approach is to manipulate the binding of the substrate to have better affinity than the actual substrate by modifying the charges in the substrate residues thereby increasing the number of interaction points. This is only possible when we have the complex structures available for each of the serotypes with their substrates. However, the idea will be to modify the substrate in such a way that it has the minimal length requirement but with better affinity than the actual substrate and also not being cleaved by the enzyme.

Scientists have attempted to crystallize native enzyme with substrates without success and that prompted the complex formation with a mutant (*13*, *23*). It points to the fact that it is difficult to obtain a complex of the enzyme with cleaved product(s). If there are many interactions between the substrate and the enzyme, it is not clear why cleavage of one bond can let the product move away from the enzyme. Either the cleaved products could change their conformation and lose interactions with the enzyme or uncleaved substrates compete with cleaved products to displace them. In view of this, if an effective inhibitor is to be developed, it needs to have stronger interactions with the enzyme than substrate but not necessarily similar to what is involved in substrate recognition. An understanding of the molecular mechanism of the changes leading to the separation of the cleaved substrate from the enzyme would be helpful in inhibitor design.

ACKNOWLEDGMENT

We thank SGX-CAT (APS) and NSLS for providing beam time for X-ray data collection and J. Romeo for technical assistance. We also thank Dr. Kumaran for his help in data collection at the APS.

REFERENCES

1. Schiavo, G., Rossetto, O., and Montecucco, C. (1994) Clostridial neurotoxins as tools to investigate the molecular events of neurotransmitter release, *Semin. Cell Biol. 5*, 221−229.
2. Montecucco, C., Papini, E., and Schiavo, G. (1994) Bacterial protein toxins penetrate cells via a four-step mechanism, *FEBS Lett. 346*, 92−98.
3. Harvey, S. M., Sturgeon, J., and Dassey, D. E. (2002) Botulism due to *Clostridium baratii* type F toxin, *J. Clin. Micrbiol. 40*, 2260−2262.
4. McCroskey, L. M., Hatheway, C. L., Woodruff, B. A., Greenberg, J. A., and Jurgenson, P. (1991) Type F botulism due to neurotoxigenic *Clostridium baratii* from an unknown source in an adult, *J. Clin. Microbiol. 29*, 2618−2620.
5. Paisley, J. W., Lauer, B. A., and Arnon, S. S. (1995) A second case of infant botulism type F caused by *Clostridium baratii*, *Pediatr. Infect. Dis. J. 14*, 912−914.
6. Humeau, Y., Dousseau, F., Grant, N. J., and Poulain, B. (2000) How botulism and tetanus neurotoxins block neurotransmitter release, *Biochimie 82*, 427−446.
7. Schiavo, G., Matteoli, M., and Montecucco, C. (2000) Neurotoxins affecting neuroexocytosis, *Physiol. Rev. 80*, 717−766.

8. Bennett, M. K., and Scheller, R. H. (1994) Molecular correlates of synaptic vesicle docking and fusion, *Curr. Opin. Neurobiol. 4*, 324−329.
9. Rossetto, O., Gorza, L., Schiavo, G., Schiavo, N., Scheller, R. H., and Montecucco, C. (1996) VAMP/synaptobrevin isoforms 1 and 2 are widely and differentially expressed in nonneuronal tissues, *J. Cell Biol. 132*, 167−179.
10. Rossetto, O., Schiavo, G., Montecucco, C., Poulain, B., Deloye, F., Lozzi, L., and Shone, C. C. (1994) SNARE motif and neurotoxins, *Nature 372*, 415−416.
11. Pellizzari, R., Rossetto, O., Lozzi, L., Giovedí, S., Johnson, E., Shone, C. C., and Montecucco, C. (1996) Structural determinants of the specificity for synaptic vesicle-associated membrane protein/synaptobrevin of tetanus and botulinum type B and G neurotoxins, *J. Biol. Chem. 271*, 20353−20358.
12. Pellizzari, R., Mason, S., Shone, C. C., and Montecucco, C. (1997) The interaction of synaptic vesicle-associated membrane protein/synaptobrevin with botulinum neurotoxins D and F, *FEBS Lett. 409*, 339−342.
13. Breidenbach, M. A., and Brunger, A. (2004) Substrate recognition strategy for botulinum neurotoxin serotype A, *Nature 432*, 925−929.
14. Hanson, M. A., and Stevens, R. C. (2000) Cocrystal structure of synaptobrevin-ll bound to botulinum neurotoxin type B at 2.0 Å resolution, *Nat. Struct. Biol. 7*, 687−692.
15. Matthews, B. W. (1988) Structural basis of the action of thermolysin and related zinc peptides, *Acc. Chem. Res. 21*, 333−340.
16. Agarwal, R., Binz, T., and Swaminathan, S. (2005) Analysis of active site residues of botulinum neurotoxin E by mutational, functional and structural studies: Glu335Gln is an apoenzyme, *Biochemistry 44*, 8291−8302.
17. Agarwal, R., Eswaramoorthy, S., Kumaran, D., Dunn, J. J., and Swaminathan, S. (2004) Cloning, high level expression, purification and crystallization of the full length Clostridium botulinum neurotoxin type E light chain, *Protein Expression Purif. 34*, 95−102.
18. Vagin, A. A., and Isupov, M. N. (2001) Spherically averaged phased translation function and its application to the search for molecules and fragments in electron density map, *Acta Crystallogr. D57*, 1451−1456.
19. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
20. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement, *Nat. Struct. Biol. 6*, 458−463.
21. Jones, T. A., Zou, J., Cowtan, S., and Kjeldgaard, M. (1991) Improved methods in building protein models in electron density map and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.
22. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszwewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Somonsom, T., and Warren, G. L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
23. Agarwal, R., Eswaramoorthy, S., Kumaran, D., Binz, T., and Swaminathan, S. (2004) Structural analysis of botulinum neurotoxin type E catalytic domain and its mutant Glu212→Gln reveals the pivotal role of the Glu212 carboxylate in the catalytic pathway, *Biochemistry 43*, 6637−6644.
24. Segelke, B., Knapp, M., Kadkhodayan, S., Balhorn, R., and Rupp, B. (2004) Crystal structure of *Clostridium botulinum* neurotoxin protease in a product-bound state: Evidence for noncanonical zinc protease activity, *Proc. Natl. Acad. Sci. U.S.A. 101*, 6888−6893.
25. Ahmed, S. A., Byrne, M. P., Jensen, M., Hines, H. B., Brueggemann, E., and Smith, L. A. (2001) Enzymatic autocatalysis of botulinum A neurotoxin light chain, *J. Protein Chem. 20*, 221−231.
26. Rao, K. N., Kumaran, D., Binz, T., and Swaminathan, S. (2005) Structural studies on the catalytic domain of clostridial tetanus toxin, *Toxicon 45*, 929−939.
27. Binz, T., Bade, S., Rummel, A., Kollewe, A., and Alves, J. (2002) Arg^{362} and Tyr^{365} of the botulinum neurotoxin type A light chain are involved in transition state stabilization, *Biochemistry 41*, 1717−1723.
28. Rossetto, O., Caccin, P., Rigoni, M., Tonello, F., Bortoletto, N., Stevens, R. C., and Montecucco., C. (2001) Active-site mutagenesis of tetanus neurotoxin implicates TYR-375 and GLU-271 in metalloproteolytic activity, *Toxicon 39*, 115−1159.
29. Rigoni, M., Caccin, P., Johnson, E. A., Montecucco, C., and Rossetto, O. (2001) Site-directed mutagenesis identifies active-site residues of the light chain of botulinum neurotoxin type A, *Biochem. Biophys. Res. Commun. 288*, 1231−1237.
30. Li, L., Binz, T., Niemann, H., and Singh, B. R. (2000) Probing the mechanistic role of glutamate residues in the zinc-binding motif of type A botulinum neurotoxin light chain, *Biochemistry 39*, 2399−2405.
31. Swaminathan, S., Eswaramoorthy, S., and Kumaran, D. (2004) Structure and enzymatic activity of botulinum neurotoxins, *Movement Disord. 19* (Suppl. 8), S17−S22.
32. Schmidt, J. J., and Stafford, R. G. (2005) Botulinum neurotoxin serotype F: identification of substrate recognition requirements and development of inhibitors with low nanomolar affinity, *Biochemistry 44*, 4067−4073.
33. Lacy, D. B., and Stevens, R. C. (1999) Sequence homology and structural analysis of clostridial neurotoxins, *J. Mol. Biol. 291*, 1091−1104.
34. Binz, T., Kurazono, H., Popoff, M. R., Eklund, M. W., Sakaguchi, G., Kozaki, S., Krieglstein, K., Henschen, A., Gill, D. M., and Niemann, H. (1990) Nucleotide sequence of the gene encoding Clostridium botulinum neurotoxin type D, *Nucleic Acids Res. 18*, 5556.
35. Shone, V. C., and Roberts, A. K. (1994) Peptide substrate specificity and properties of the zinc-endopeptidase activity of botulinum type B neurotoxin, *Eur. J. Biochem. 225*, 263−270.
36. Schmidt, J. J., and Bostian, K. A. (1997) Endoproteinase activity of typeA botulinum neurotoxin:substrate requirements and activation by serumalbumin, *J. Protein Chem. 16*, 19−26.
37. Vaidyanathan, V. V., Yoshino, K., Jahnz, M., Dorries, C., Bade, S., Nauenburg, S., Niemann, H., and Binz, T. (1999) Proteolysis of SNAP-25 isoforms by botulinum neurotoxin types A, C, and E: Domains and amino acid residues controlling the formation of enzyme-substrate complexes and cleavage, *J. Neurochem. 72*, 327−337.
38. Simpson, L. L., Maksymowych, A. B., Park, J.-B., and Bora, R. S. (2004) The role of the interchain disulfide bond in governing the pharmacological actions of botulinum toxin, *J. Pharmacol. Exp. Ther. 308*, 857−864.
39. Ahmed, S. A., and Smith, L. A. (2000) Light chain of botulinum A neurotoxin expressed as an inclusion body from a synthetic gene is catalytically and functionally active, *J. Protein Chem. 19*, 475−487.
40. Fuchs, O., Babusiak, M., Vyoral, D., and Petrak, J. (2003) Role of zinc in eukaryotic cells, zinc transporters and zinc-containing proteins, *Sb. Lek. 104*, 157−170.
41. Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of proteins, *J. Appl. Crystallogr. 24*, 946−950.
42. DeLano, W. L. (2002) in *The PyMOL User's Manual*, DeLano Scientific, San Carlos, CA.

BI0510072

Appendix 4

ELSEVIER

Toxicon 45 (2005) 929–939

TOXICON

www.elsevier.com/locate/toxicon

Structural analysis of the catalytic domain of tetanus neurotoxin

Krishnamurthy N. Rao[a], Desigan Kumaran[a], Thomas Binz[b], Subramanyam Swaminathan[a,*]

[a]*Brookhaven National Laboratory, Biology Department, 50 Bell Avenue, Upton, NY 11973, USA*
[b]*Department of Biochemistry, Medizinischule Hochschule Hannover, Hannover, Germany*

Received 10 December 2004; revised 16 February 2005; accepted 17 February 2005
Available online 13 April 2005

Abstract

Clostridium neurotoxins, comprising the tetanus neurotoxin and the seven antigenically distinct botulinum neurotoxins (BoNT/A-G), are among the known most potent bacterial protein toxins to humans. Although they have similar function, sequences and three-dimensional structures, the substrate specificity and the selectivity of peptide bond cleavage are different and unique. Tetanus and botulinum type B neurotoxins enzymatically cleave the same substrate, vesicle-associated membrane protein, at the same peptide bond though the optimum length of substrate peptide required for cleavage by them is different. Here, we present the first experimentally determined three-dimensional structure of the catalytic domain of tetanus neurotoxin and analyze its active site. The structure provides insight into the active site of tetanus toxin's proteolytic activity and the importance of the nucleophilic water and the role of the zinc ion. The probable reason for different modes of binding of vesicle-associated membrane protein to botulinum neurotoxin type B and the tetanus toxin is discussed. The structure provides a basis for designing a novel recombinant vaccine or structure-based drugs for tetanus.
Published by Elsevier Ltd.

Keywords: *Clostridium* neurotoxin; Tetanus neurotoxin; Botulinum neurotoxin; Zinc; Metalloprotease; Dual-wavelength anomalous diffraction (DAD); X-ray structure

1. Introduction

Tetanus neurotoxin (TeNT) produced by *Clostridium tetani* and the seven antigenically distinct botulinum neurotoxins (BoNT/A-G) produced by *Clostridium botulinum* together constitute the family of clostridial neurotoxins (CNTs) (Montecucco and Schiavo, 1995; Rawlings and Barrett, 1995). Tetanus neurotoxin acts on the central nervous system and inhibits the release of glycine and γ-aminobutyric acid causing spastic paralysis, tetanus (Galazka and Gasse, 1995). In contrast, botulinum neurotoxins (BoNTs) act on the peripheral nervous system and inhibit the release of acetylcholine at the neuromuscular junction, causing flaccid paralysis, botulism (Simpson, 1986). These proteins are of public health concern as paralysis by these toxins still takes hundreds of lives every year (Galazka and Gasse, 1995). They are also emerging as biowarfare threats.

CNTs consist of three functional domains: binding, translocation, and catalytic. CNTs bind to the neuronal cells via gangliosides and a second protein receptor and then are internalized. Their catalytic domain is translocated through the vesicle membrane into the cytosol where it attacks and cleaves one of the proteins forming the core of the synaptic vesicle fusion apparatus. CNTs synthesized as 150 kDa inactive single chain molecules are post-translationally modified either by endogenous or exogenous proteases into two polypeptide chains covalently linked by a disulfide bond: the C-terminal heavy chain (HC, 100 kDa)

* Corresponding author. Tel.: +1 631 344 3187; fax: +1 631 344 3407.
E-mail address: swami@bnl.gov (S. Swaminathan).

0041-0101/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.toxicon.2005.02.032

and the N-terminal light chain (LC, 50 kDa). The C-terminal heavy chain is responsible for the binding of the toxin to specific neuronal receptors and the translocation of the N-terminal LC, the catalytic domain, into the neuronal cytosol. The LCs of CNTs contain the zinc-binding motif HExxH + E and are accordingly zinc-dependent metalloproteinases causing toxicity (Eswaramoorthy et al., 2004; Fillippis et al., 1995; Kurazono et al., 1992; Schiavo et al., 1994; Simpson et al., 2001). Interestingly, though CNTs display high sequence homology, have similar functions and probably have similar structure, they are unique in that their target protein and the scissile bonds they cleave are different, a property very commonly shared by members of a given protease family. TeNT, BoNT/B, /D, /F, and /G cleave vesicle-associated membrane protein (VAMP) (Schiavo et al., 2000). While the cleavage site is different for D, F and G, it is the same for TeNT and BoNT/B, a unique case in CNTs. Similarly, BoNT/A, /E, and C cleave synaptosomal-associated protein 25 kDa (SNAP-25) at different peptide bonds (Binz et al., 1994; Schiavo et al., 1993; Vaidyanathan et al., 1999). BoNT/C is unique in that it also cleaves syntaxin.

Chemically modified TeNT is the most used human vaccine and BoNTs are increasingly used in the therapy of human diseases caused by hyperfunction of cholinergic nerve terminals (Jankovic and Hallett, 1994). Chemically treated toxoid vaccine is available against BoNTs but no therapeutic treatments exist as of now. To develop an effective structure-based vaccine/inhibitor/antitoxin to treat tetanus victims, an understanding of the molecular mechanism at the atomic level, especially the peptidase activity by the catalytic domain, is a prerequisite. Though experimental three-dimensional structures are available for some of the CNTs and their functional fragments, no experimental structure is available for TeNT-LC except for a homology-based model (Agarwal et al., 2004a; Breidenbach and Brunger, 2004; Emsley et al., 2000; Hanson and Stevens, 2000; Lacy and Stevens, 1998; Rossetto et al., 2001; Segelke et al., 2004; Swaminathan and Eswaramoorthy, 2000; Umland et al., 1997). These structures helped to map the active site and the residues forming it. Although the active sites are similar, the specificity and selectivity of CNTs suggest that there must be additional factors that define the substrate specificity. Thus, it becomes important to understand the molecular structure of each of the CNTs at the atomic level to gain insight into their unique ability to cleave at specific scissile bonds within the same substrate and also among different substrates. A possible catalytic mechanism has been proposed for TeNT from the homology model (Rossetto et al., 2001). Extensive mutational studies have been carried out on TeNT (Li et al., 1994; McMahon et al., 1993; Yamasaki et al., 1994). Some of these mutants were proteolytically inactive indicating that they may have a direct role in the catalytic activity of the protease. However, there was no matching X-ray structural information available to interpret these results. The chemical characterization of the active site and of the residues involved in zinc binding of TeNT is important to understand the mechanism of the proteolytic activity of this novel family of metalloproteases (Montecucco and Schiavo, 1995; Schiavo et al., 1994) and may lead to the development of novel and safer recombinant vaccine produced outside of *C. tetani*. A crystal structure analysis provides a tremendous amount of insight into both the structure and the function of the protein. Although the crystallization of TeNT-LC has been reported, no experimental structure is available in literature (Tonello et al., 1994). Here, we present the first experimental three-dimensional structure of TeNT-LC and compare it with the crystal structures of available BoNT-LC structures.

2. Materials and methods

2.1. Expression and purification of his-tagged TeNT-LC protein

The procedure for expression and purification is as described in Agarwal et al. (2004b) except that ampicillin-containing medium was used. The protein was eluted from Ni-NTA agarose columns with increasing concentrations of imidazole buffer. Fractions of the eluate were analyzed by SDS-PAGE using 4–20% gels followed by staining with Coomassie blue. A ~52 kDa band corresponding to TeNT-LC reproducibly eluted in 50–100 mM imidazole fractions was obtained. Recovery of TeNT-LC was more than 4 mg/l of induced cell culture. At this stage, it is nearly 80% pure. TeNT-LC was further purified by gel filtration on a (2×20 in.) column of Superdex-75 using Akta FPLC which also helped in exchanging the buffer from phosphate to 20 mM HEPES, pH 7.2 + 200 mM NaCl. Peak fractions containing only pure TeNT-LC were pooled and concentrated to ~10.0 mg/ml using Centriprep YM-10.

2.2. Enzymatic activity of TeNT-LC

The proteolytic activity of TeNT-LC was assayed in vitro on its substrate VAMP which had an N-terminal GST tag. The assay was performed in a final volume of 20 μL [20 mM HEPES, pH 7.4, 2 mM DTT, 10 μM $Zn(CH_3COO)_2$] containing a 10–5000 nM concentration of LC and a 5 μM concentration of VAMP and the mixtures were incubated at 37 °C for 30 min. The reactions were stopped by adding 10 μl of 3× concentrated SDS-PAGE sample buffer. The extent of cleavage was then evaluated following SDS-PAGE (4–20% gels) by the appearance and intensity of a new band at ~30 kDa due to the cleavage of GST–VAMP at Gln76–Phe77 peptide bond of VAMP. A 1000 nM concentration of LC was required to cleave 50% of the substrate in 30 min.

2.3. Crystallization

The crystallization screening was carried out by the sitting drop vapor diffusion method using Hampton Research high throughput crystallization screens and TECAN Genesis Freedom crystallization robot. One microlitre of the protein solution containing 10 mg/ml protein, 200 mM NaCl, and 20 mM HEPES buffer at pH 7.2 was combined with 1 μl of the precipitant and equilibrated against 600 μl of the same precipitant. Diffraction-quality crystals obtained at room temperature using 20% PEG 4000, 0.2 M $MgCl_2$, and 0.1 M HEPES at pH 7.0 as precipitant grew to their full size in about 4 weeks. Mother liquor containing 15–20% glycerol proved to be a good cryoprotectant for collecting X-ray diffraction data at liquid nitrogen temperature. Crystal data are presented in Table 1.

2.4. Data collection

Data were collected at liquid nitrogen temperature at the National Synchrotron Light Source, Brookhaven National Laboratory. Since the protein contains one zinc ion per molecule, dual-wavelength anomalous dispersion (DAD) data were collected from a wild-type LC crystal at wavelengths 1.2810 and 1.2824 Å, corresponding to the peak and inflection of the zinc absorption edge at beamline X12C, NSLS with a CCD-based Brandeis B4 detector. Data covering a total of 360° rotation in ϕ were collected for an oscillation range of 1° per frame. However, the diffraction limit extended to 3.0 Å only. Higher resolution data to 2.4 Å resolution were later collected at beamline X25, NSLS. Data were processed with HKL2000 (Otwinowski and Minor, 1997). Details of data collection statistics are given in Table 1.

2.5. Structure determination

The peak and edge data were used to locate the position of zinc atom using SOLVE (Terwilliger and Berendzen, 1997). Initial model phases were also calculated by the molecular replacement method using AmoRe with BoNT/B-LC (PDB idcode: 1F82) as search model (Navaza and Saludjian, 1997). AmoRe gave a good starting model with *R* and correlation coefficient of 0.47 and 0.58, respectively. Phases obtained with zinc were refined by SHARP (De-La-Fortelle and Bricogne, 1997) using the model phases as external phases and further improved by density modification (Cowtan, 1994). The molecular model with the sequence of TeNT-LC was built using the program O (Jones et al., 1991). At this stage, since higher resolution data were available, they were used in the refinement.

Table 1
Data collection and refinement statistics

	Crystal-I		Crystal-II
Space group	*C222*		*C222*
Cell parameters	$a=105.49$ Å, $b=176.92$ Å, and $c=57.45$ Å		$a=106.61$ Å, $b=177.29$ Å, and $c=54.55$ Å
Phasing			
Method: DAD	Peak	Edge	Native
Wavelength (Å)	1.2810	1.2824	1.10
Anomalous scatterer	Zinc	Zinc	
Resolution (Å)	50–3.0	50–3.0	50–2.4
No. of reflections	15,350	15,365	20,639
Rmerge[a] (%)	0.13 (0.35)	0.13 (0.35)	0.12 (0.39)
Overall completeness (%)	99.8 (99.3)	98.5 (98.1)	99.4 (99.1)
Redundancy	13.3 (9.7)	12.8 (9.3)	9.4 (9.3)
Overall FOM	0.20		
After solvent flattening	0.90		
Refinement statistics			
Resolution range (Å)	50–2.6		
No. of reflections	15,170		
Overall completeness (%)	92.7 (90)		
R-factor	0.21		
R-free	0.27		
No. of protein atoms	3221		
No. of heterogen atoms	1		
No. of water molecules	113		
RMS deviations			
Bond lengths (Å)	0.007		
Bond angles (°)	1.33		

[a] The values corresponding to the outermost shell are given within parentheses. Rmerge is as defined in HKL2000. FOM, R-factor and R-free are as defined in SHARP program.

After initial rigid body refinement, the structure was further refined using CNS (Brunger et al., 1998). Solvent molecules were added from the difference Fourier map. The protein model is complete except for residues 64–67, 208–219, 252–264 which are in the loop regions and for about 30 C-terminal residues, for which the electron density is poorly defined, indicating these regions may be disordered. The final *R* and *R*-free values are 0.21 and 0.27, respectively. The structure has been validated with PROCHECK (Laskowski et al., 1993), and the distribution of residues in the Ramachandran plot is given in Table 1. Coordinates have been deposited with the Protein Data Bank (1YVG).

3. Results

3.1. Structure of TeNT-LC

TeNT-LC forms a dimer in the crystal via a crystallographic two-fold with about 10% (3100 Å^2) of the total surface area buried at the interface. A Ribbons representation of the catalytic domain dimer of TeNT-LC along with the active site residues is shown in Fig. 1a. All of the LC structures so far determined, except BoNT/B-LC, exist as dimer either via crystallographic or non-crystallographic two-fold. While the dimeric interface covers the active site

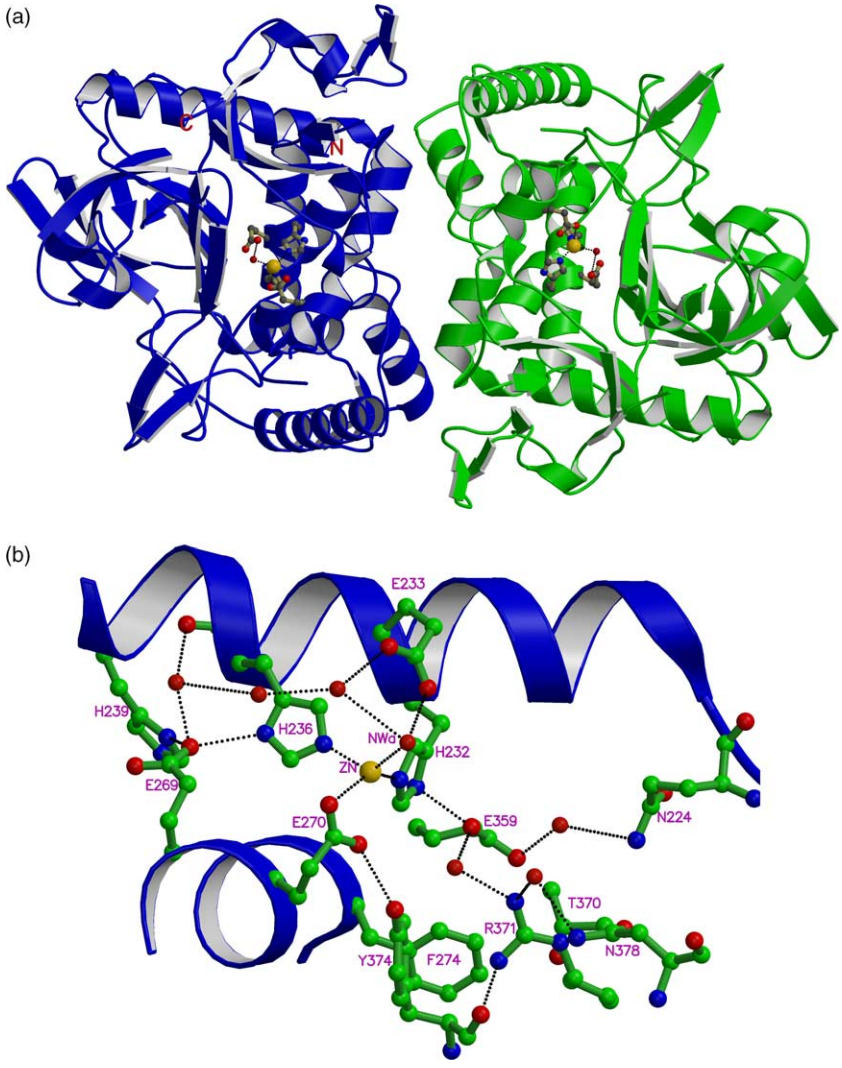

Fig. 1. (a) Ribbons representation of a dimer of TeNT-LC, with zinc and the coordinating residues shown in the ball-and-stick model. The dimer is formed by a crystallographic two-fold symmetry passing through the mid point normal to the plane of the figure. (b) The active site of TeNT-LC showing the interactions and hydrogen bonding network involving the residues that are nearly conserved in TeNT, BoNT/B, /E and /A LCs. Hydrogen bonds are shown as dashed lines in black. The nucleophilic water molecule coordinated with zinc is labeled 'NWa'.

zinc in BoNT/A-LC, the active sites are exposed to the solvent region in BoNT/E and TeNT-LCs.

As in other LC structures, the active site of TeNT-LC is located deep inside a cavity by which the substrate gains access to the active site. The active site is centered around a zinc cation coordinated by a strictly conserved HExxH + E motif. In TeNT-LC, zinc is directly coordinated by residues His232, His236 and Glu270 (the numbering scheme is as in Swiss-Prot: P04958) (Fig. 1b). The fourth ligand is a nucleophilic water molecule that also forms a strong hydrogen bond with Glu233. The coordinating distances are given in Table 2. The coordination geometry at zinc is slightly distorted from tetrahedral geometry. The δ-carboxyl group of Glu233 serves to coordinate the hydrogen-bonded nucleophilic water molecule in one of the four tetrahedral positions around the catalytic zinc ion. This interaction seems to be important for the activation of the nucleophilic water. In most zinc endopeptidases, the substrate or an inhibitor displaces the nucleophilic water and moves it closer to the glutamate residue serving as proton shuttling agent. Because of this movement, the nucleophilic water would make stronger hydrogen bonds (possibly two) with the proton shuttling glutamate, making it a strong nucleophile (Swaminathan et al., 2004). The residues coordinated with the zinc form a primary sphere of residues centered around the active site and their direct role in catalysis has been confirmed through site-directed mutagenesis studies (Li et al., 1994; Rossetto et al., 2001; Yamasaki et al., 1994). From mutagenesis experiments, it also appears that residues in the secondary layer near the active site are essential or play a crucial role in the catalytic function. These residues found in the vicinity (within 10 Å from zinc) of the active site zinc include Glu233, His239, Phe274, Glu359, Thr370, Arg371, Tyr374, and Asn378. Among these, mutation of Glu233, His239, Phe274, Arg371 and Tyr374 residues in TeNT and the corresponding residues in BoNTs resulted in complete or partial loss of the catalytic activity indicating their direct role in catalytic function. These residues are found farther than 3.0 Å from the zinc or the nucleophilic water and their interacting distances are given in Table 2. These along with a few water molecules form hydrogen bonds to the residues coordinated with the zinc ion (Fig. 1b), indicating that these residues are essential for the stability of the structure and/or the conformation of the active site.

Table 2
Active site residues interacting distances (Å) in TeNT, BoNT/B, and /E LC structures

	TeNT-LC (Å)	BoNT/ B-LC (Å)	BoNT/ E-LC (Å)
Zn–H232 NE2	2.37	2.11	2.18
Zn–H236 NE2	2.13	2.15	2.16
Zn–E270 OE1	2.49	2.60	2.22
Zn–E270 OE2	2.46	2.20	2.46
Zn–Nu. water	2.50	2.08	2.17
Nu. water–E233OE1	3.84	3.54	4.03
Nu. water–E233OE2	3.09	2.80	2.82
Nu. water–Y374–OH	6.06	4.62	3.56
Nu. water–E270 OE1	3.62	3.11	3.15
Nu. water–R371NH2	7.90	7.46	7.06
E359OE1–H232ND1	2.65	2.71	2.61
E359OE2–R371NH1	4.42	3.04	2.99
E269OE1–H236ND1	2.79	3.03	2.76
E269OE1–H239ND1	2.58	2.73	2.73
Y374–OH–Zn	5.37	4.30	3.94
Y374–OH–E270OE1	3.10	3.69	3.15
Y374–OH–E270OE2	4.48	3.70	3.43
R371NH–Zn	7.29	6.50	7.08

The numbering scheme corresponds to TeNT.

The TeNT-LC model excludes the 60 (residues 64–67), 200 (residues 208–219) and 250 (residues 252–264) loops since the electron density is poorly defined in these regions. Residues in these loop regions, in general, appear disordered in all LC structures. A probable reason is that they lose hydrophobic interactions with the translocation domain that were present in the holotoxin on separation or lack of it in the recombinant LC. The C-terminal region which forms the central strand of a β sheet with the two outer strands from the translocation domain in the holotoxin may similarly be disordered. These catalytic domain loops were observed to change their conformation significantly upon separation from the holotoxin and subsequent binding of the substrate in the case of BoNT/B (Hanson and Stevens, 2000). In fact, the substrate is found to bind between these loops. Thus, efficient proteolysis of the substrate seems dependent on the movement of these loops because it exposes the binding surface of the catalytic domain and brings the catalytic residues in position and facilitates the catalysis. Unfortunately, the poorly defined electron density prevents a detailed comparison of these loops with those of substrate bound and unbound BoNT/B-LC. However, when the model of TeNT-LC is superposed on native or substrate bound BoNT/B-LC, the weak electron density at least allows tracing the direction of these loops. Thus, it is very likely that these loops in TeNT-LC undergo similar rearrangement upon substrate binding as in the case of BoNT/B-LC and, in addition, this cooperative change is supported based on kinetic data on TeNT (Cornille et al., 1997). Comparison with BoNT/B-LC and the packing consideration of the *C222* space group suggest that the orientation of these loops may be different from BoNT/B-LC as otherwise there will be steric interactions. The position of loop 140 (residues 140–148) in TeNT-LC looks to be different than that observed in BoNT/B-LC. The loop 140 of TeNT-LC is tilted away from the corresponding BoNT/B-LC loop with the farthest distance between corresponding residues being approximately 12 Å. A possibility that the missing residues in the 250 loop may be due to auto catalysis as in the case of BoNT/A-LC was ruled out since the SDS-PAGE gel analysis showed no such indication (Segelke et al., 2004). The disorder in the C-terminal region seems to be common

to all LCs. This may also be due to the lack of the interchain disulfide bond.

3.2. Comparison with other CNT LCs

The structure of TeNT-LC is similar to those of BoNT/A, B, and E-LCs indicating that all the LCs of CNTs have similar structure and are unique only in that they have stringent substrate specificity. The specificity of CNTs is therefore likely to arise from the residues that form the channel by which the substrate gains access to the active site and the complementarity of substrate and the enzyme residues. When the C_α atoms of TeNT-LC were superimposed onto those of BoNT-LCs using the program LSQMAN the rms deviation varied between 1 and 1.4 Å (1.02 Å for 380 C_α of BoNT/B, 1.2 Å for 350 C_α of BoNT/E, 1.4 Å for 340 C_α of BoNT/A) (Kleywegt and Jones, 1997). A stereo view of the C_α trace of TeNT-LC superimposed on other BoNT-LC models is shown in Fig. 2a. The TeNT-LC structure superimposes better with BoNT/B than with BoNT/A or BoNT/E-LC. This is not surprising since TeNT and BoNT/B have the substrate specificity and scissile bond selectivity for VAMP among these CNTs. This may also be partly because TeNT and BoNT/B LCs share 52% sequence identity which is greater than that (∼35%) shared by all clostridial neurotoxin LCs (Kurazono et al., 1992; Lacy and Stevens, 1999). An alignment of sequences of LCs of TeNT, and BoNT/B is shown in Fig. 3.

When the active sites of TeNT and BoNT/B are superposed, they preserve some of the structural features involving a few conserved residues. These include Asn168, Glutamate residues 169, 233, 269, 359, His239, Thr273, Arg371, Ile235 and Phe274 and the zinc-binding motif residues HExxH (Fig. 2b). Many of the interactions established by these residues are common to all CNTs (Agarwal et al., 2004a). However, it is intriguing to note here that although TeNT and BoNT/ B process the same peptide bond in VAMP, there are some differences in details near the active site. Three catalytically important residues,

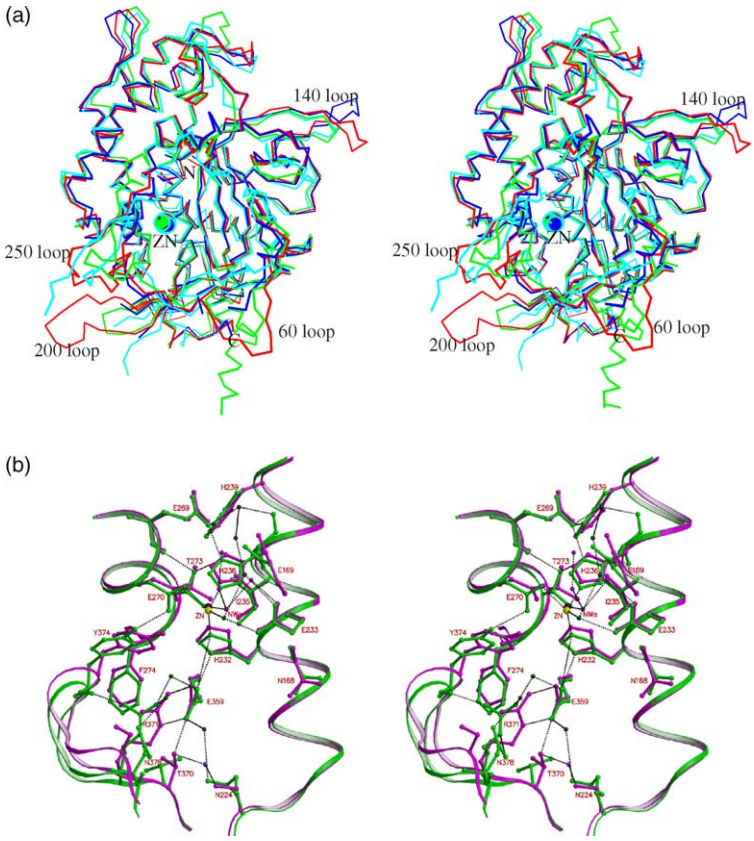

Fig. 2. (a) Stereo view of a superposition of C_α traces of TeNT (blue), BoNT/B (red), /E (green), /A (light blue). The conformations of all the LCs are similar except for the loop regions. The loops 60, 200 and 250 are incomplete in the TeNT-LC structure. The loop140 points in the opposite directions in TeNT and BoNT/B LCs. Zinc atoms are shown in different sizes for clarity. (b) Stereo view of a superposition of the active sites of TeNT (green) and BoNT/B (magenta) LCs showing the differences in the interactions of the three catalytically important residues Glu359, Arg371 and Tyr374 in the two LC structures. The residues shown are within 10 Å distance from the zinc position. The interactions involving the hydrogen bonds are shown in dashed lines. The water molecules of TeNT are shown in dark green and those of BoNT/B are shown in purple.

```
TENT            PITINNFRYSDPVNNDTIIMMEPPYCKGLDIYYKAFKITDRIWIVPERYEFGTKPEDFNP 60
BONT/B          PVTINNFNYNDPIDNNNIIMMEPPFARGTGRYYKAFKITDRIWIIPERYTFGYKPEDFNK 60
                *:*****.*.**::*:.*******:.:* . **************:**** ** ******

TENT            PSSLIEG-ASEYYDPNYLRTDSDKDRFLQTMVKLFNRIKNNVAGEALLDKIINAIPYLGN 119
BONT/B          SSGIFNRDVCEYYDPDYLNTNDKKNIFLQTMIKLFNRIKSKPLGEKLLEMIINGIPYLGD 120
                .*.:::  ..*****:**.*:..*: *****:*******.:  ** **: ***.*****:

TENT            SYSLLDKFDTNSNSVSFNLLEQDPSGATTKSAMLTNLIIFGPGPVLNKNEVRGIVLRVDN 179
BONT/B          RRVPLEEFNTNIASVTVNKLISNPGEVERKKGIFANLIIFGPGPVLNENET----IDIGI 176
                    *::*:**  **:.* * .:*. .  *..:::************:**.    : :.

TENT            KNYFPCRDGFGSIMQMAFCPEYVPTFDNVIENITSLTIGKSKYFQDPALLLMHELIHVLH 239
BONT/B          QNHFASREGFGGIMQMKFCPEYVSVFNNVQENKGASIFNRRGYFSDPALILMHELIHVLH 236
                :*:*..*:***.**** ******..*:** **  :  :.:  **.****:**********

TENT            GLYGMQVSSHEIIPSKQEIYMQHTYPISAEELFTFGGQDANLISIDIKNDLYEKTLNDYK 299
BONT/B          GLYGIKVDDLPIVPNEKKFFMQSTDAIQAEELYTFGGQDPSIITPSTDKSIYDKVLQNFR 296
                ****::*..  *:*.:::::** * .*.****:******..:*: . .:.:*:*.*:::: 

TENT            AIANKLSQVTSCN-DPNIDIDSYKQIYQQKYQFDKDSNGQYIVNEDKFQILYNSIMYGFT 358
BONT/B          GIVDRLNKVLVCISDPNININIYKNKFKDKYKFVEDSEGKYSIDVESFDKLYKSLMFGFT 356
                .*.::*.:*  *  ****:*: **: :::**:* :**:*:* :: :.*: **:*:*:*** 

TENT            EIELGKKFNIKTRLSYFSMNHDPVKIPNLLDDTIYNDTEGFNIESKDLKSEYKGQNMRVN 418
BONT/B          ETNIAENYKIKTRASYFSDSLPPVKIKNLLDNEIYTIEEGFNISDKDMEKEYRGQNKAIN 416
                * ::.::::**** **** .  **** **** **.  *****..**::.**:***  :* 

TENT            TNAFRNVDGSGLVSKLIGLCKKIIPPTNIRENLYNRT 455
BONT/B          KQAYEEISKEHLAVYKIQMCKSVKAPG---------- 443
                .:*:.::. . *.   * :**.: .*
```

Fig. 3. The alignment of sequences of TeNT and BoNT/B LCs using CLUSTAL-W. The residues seen within 4 Å distance from the substrate VAMP are highlighted. The residues coordinated to zinc are in bold inside the box. The bottom line represents the degree of homology—'*' identical, ':' conserved and '.' semiconserved.

Glu359, Arg371 and Tyr374, were found to be different with respect to their position and interactions they make near the active site (Fig. 2b). Through mutagenesis experiments on TeNT and BoNT/A-LCs and structural studies on various serotypes of BoNTs, it has been shown that these residues are involved in the catalysis (Agarwal et al., 2004a; Binz et al., 2002; Hanson and Stevens, 2000; Rossetto et al., 2001; Segelke et al., 2004). In TeNT, the OE1 atom of Glu359 forms a hydrogen bonded interaction with ND1 atom of the His232, which is coordinated with the zinc ion, and thus helps in orienting the side chain of His232. The Glu359 side chain OE1 and OE2 atoms are held in position through hydrogen bond interactions with two water molecules. These two water molecules in turn interact with NH2 of Arg371 and NE2 of Asn224. However, in BoNT/B the side chain atoms of the equivalent Glu357 directly interact with the NH2 of the equivalent Arg369 and OG1 of Thr368. The side chain of the equivalent residue Thr370 in TeNT-LC is positioned away from this site. However, the side chain of Asn224 in TeNT-LC provides an interaction. Interestingly, the position of Asn224 in TeNT-LC is occupied by Ser221 in BoNT/B-LC (Fig. 2b).

The NE, NH1 and NH2 atoms of Arg371 in TeNT interact with OD1 of Asn378, the main chain O atom of Tyr374 and a water molecule, respectively. However, NE and NH2 atoms of the equivalent Arg369 in BoNT/B interact with OE2 and OE1 atoms of Glu357, whereas the NH1 atom of Arg369 is too far away to interact with the main chain O atom of Tyr372. As a result, in BoNT/B-LC the Arg369 is close to Glu357 but is away from the Tyr372. In contrast, in TeNT the equivalent Arg371 moves away from the Glu359 but is close to Tyr374 and one more aromatic residue Phe274. Recently, mutation of the equivalent Phe266 in BoNT/A-LC was shown to have an effect on the catalytic activity of the toxin (Rigoni et al., 2001). The side chain of Arg371 in TeNT is positioned in between the aromatic planes of the Tyr374 and Phe274, so that the guanidino group of Arg371 interacts with the π electron clouds of the two aromatic rings to neutralize their charge (Fig. 2b). These amino–aromatic electrostatic interactions were observed to occur in the range of 3.6–6 Å (max at 4.7 Å) distance in proteins (Burley and Petsko, 1986). The NH1 atom of Arg371 was found to be at a distance of 4.18 and 3.95 Å from the aromatic planes of Tyr374 and Phe274, respectively. However, in BoNT/B-LC, these electrostatic interactions seem to be stabilized through interaction between Arg369 and Glu357.

The OH-group of Tyr374 in TeNT forms a hydrogen bond with Glu270 OE1, which in turn is coordinated with

the zinc ion (Fig. 2b). This interaction seems to be important to hold the side chain of the residue Glu270 in position. The corresponding residue in BoNT/B (Tyr372) has been implicated to be involved in stabilizing the cleaved product of the substrate by donating a proton to the leaving group, Phe77, of the VAMP (Hanson and Stevens, 2000). This hydrogen bond mediated interaction is absent in BoNT/B-LC. However, the side chain of the residue Tyr372 in BoNT/B is at a distance of 4.3 Å from the catalytic zinc ion, but the equivalent distance in TeNT is 5.37 Å. It is not clear whether this has an effect on the catalytic activity.

3.3. Differential mode of VAMP binding to TeNT and BoNT/B-LCs

Both TeNT and BoNT/B cleave VAMP at the same peptide bond, Gln76–Phe77. Nevertheless, appreciable divergence in their sequence (52% similar) reflects some differences between the two enzymes. Indeed, the two enzymes were observed to be different in their proteolytic activities towards different minimum sized VAMP substrates, pH and temperature profiles, and sensitivity to inhibitors such as captopril. It was suggested that these differences could be a consequence of subtle differences in their intracellular actions and due to fine structural details at the atomic level (Foran et al., 1994). It has been shown that BoNT/B cleaves a peptide corresponding to the VAMP segment encompassing residues 55–94 while TeNT requires residues 33–94, an amino-terminal extension of 22 residues, to cleave efficiently (Cornille et al., 1997; Foran et al., 1994; Pellizzari et al., 1996). The minimum length of substrate peptide for BoNT/B contains only the second (V2) of the two SNARE Secondary Recognition (SSR) motifs; but for TeNT the optimum substrate length contains both SSR motifs of VAMP, V1 and V2 (38–47 and 62–71). In general, it is the residues spanning the cleavage site in a substrate, which define the efficiency of a protease. Since these toxins require different minimal substrate length, they must recognize their substrate in a unique but different fashion. Studies based on mutation of VAMP residues, and antibodies raised against motif peptides or recombinant proteins have suggested that TeNT and BoNT/B were similar in their recognition of SSR motifs (Pellizzari et al., 1996). With the currently determined three-dimensional structure of TeNT-LC, it is clear that the LCs of TeNT and BoNT/B fold very similarly and their active site architecture exhibits extensive similarity although with some differences in H-bonding interactions. Hence, the difference in minimum length peptide required for activity cannot be explained by the sole recognition of the cleavage site. This might suggest that the binding interactions required for positioning the substrate for cleavage may be different. Taking together these observations, it is very likely that TeNT and BoNT/B bind VAMP at the same site but their mode of binding may differ. Thus the two toxins are likely to have extended substrate-binding regions in which a large number of amino acids interact with moieties away from the active site. Indeed, in the crystal structure of BoNT/B bound substrate product, the substrate and the protease have numerous interactions remote from the active site (Hanson and Stevens, 2000).

An extensive analysis in terms of three-dimensional structure to pinpoint the differences that may be responsible for unique recognition of VAMP by TeNT and BoNT/B LCs was carried out. First, the TeNT-LC structure was superposed on BoNT/B-LC to be in the same orientation. Then the VAMP peptide was positioned similar to its orientation in BoNT/B-LC (Hanson and Stevens, 2000). The amino acid residues that are interacting within 4 Å distance from VAMP residues that are conserved in space were compared (Fig. 4). About 50 residues that fall within 4 Å from the substrate are highlighted in Fig. 3. This excluded the residues that are within coordinating distance from the zinc ion, because they process the identical scissile bond. These residues did not differ with respect to their C_α positions in the two toxins as their position remains the same when compared to the unbound BoNT/B structure. The comparison revealed that only 30% of these residues are conserved. Forty percent of the residues differed between TeNT and BoNT/B LCs with respect to their hydrophilic or hydrophobic nature. About 30% of the residues in TeNT had bulkier side chains when compared to the residues in BoNT/B. The 30% conserved residues may have common interactions with the substrate in both the toxins. Thus, these conserved residues can be thought of as the minimal essential residues to recognize the substrate by the two toxins. The 40% residues that provide different environment with respect to their hydrophobic or hydrophilic nature around the substrate may be responsible for the differential mode of substrate binding by TeNT and BoNT/B. The 30% residues of TeNT that are bulkier (shown in yellow, Fig. 4a) compared to those in BoNT/B make steric clashes with the VAMP residues, if it were to bind as in BoNT/B. These residues would destabilize the binding of substrate by the protease. The observation that the TeNT requires extra 22 N-terminal residues for optimal activity compared to BoNT/B can be explained as follows. TeNT-LC needs to interact with a more extended segment of the substrate compared to BoNT/B-LC to stabilize the enzyme–substrate complex. The affinity to the substrate segment shared with BoNT/B could be lower. This fact is evident from comparison of the TeNT-LC structure with that of the BoNT/B structure bound to VAMP. Out of the 40 residues of VAMP bound to BoNT/B, 13 would make steric clashes in TeNT. Of this 13, five residues (Glu62, Asp64, Asp65, Asp68, and Gln71) of the substrate are in the V2 region. For example, Met377 of TeNT would produce a very short contact with Asp 68 of VAMP. This might not allow the V2 region to come close to the enzyme for binding in contrast to BoNT/B. This could be compensated by the binding of V1 to the enzyme. There are two possibilities. The V2 region and the adjacent N terminal region of the substrate could project out like a hair pin

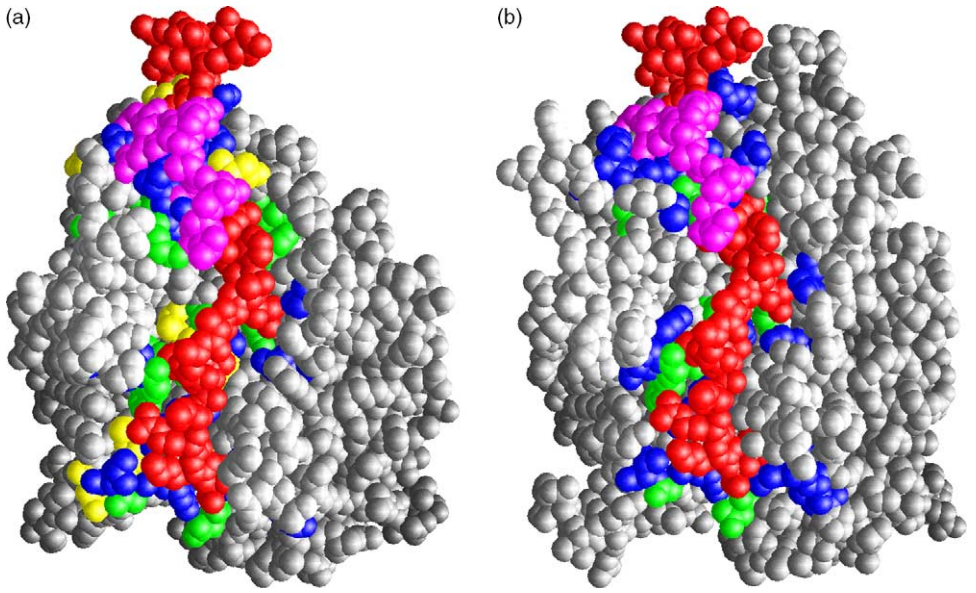

Fig. 4. (a) Sphere model representation of TeNT-LC with the substrate VAMP peptide placed in a similar position and orientation as in the BoNT/B–VAMP complex crystal structure (Hanson and Stevens, 2000). (b) Sphere model representation of BoNT/B bound to peptide VAMP. VAMP residues are shown in red except the V2 region, which is in magenta. Residues in TeNT-LC and BoNT/B-LC within 4 Å of the VAMP peptide are shown in green for conserved residues and blue if they are different. Residues which are bulkier in TeNT compared to BoNT/B are shown in yellow. The apparent difference in the shapes of TeNT-LC and BoNT/B-LC is due to the absence of some loops in TeNT-LC.

allowing the V1 region to bind in the same place as V2 in BoNT/B or the steric clash would push V2 away from the surface and V1 could bind to the enzyme in a shallow region on the other side of the molecule to increase the binding constant. However, since Asp64 and Asp68 of V2 are conserved in V1, and Asp65 of V2 is replaced by Glu41 in V1, we prefer the second possibility. These collective observations suggest that TeNT and BoNT/B interact differently with VAMP and provide a base for the design of effective inhibitors for both toxins. This situation could be similar to what was shown for BoNT/A or /E where the absence of one of the SSR motif is compensated by one of the other three SSRs (Washbourne et al., 1997).

4. Conclusion

Tetanus and botulinum neurotoxins are solely responsible for the neuroparalytic syndromes of tetanus and botulism. Considering the importance of tetanus neurotoxin in human pathology and therapeutics, there is an immediate demand for three-dimensional structural information, especially of the catalytic domain. The present experimental structure fulfils this need. Currently, anti-tetanus vaccine is prepared by chemical modification of toxin with formaldehyde. However, it is believed that genetically modified toxins better preserve immunogenicity than chemical modification and such a vaccine based on a mutated bacterial toxin is already in use as anti-pertussis vaccine (Rappuoli et al., 1992). The present crystal structure study provides a basis for developing a novel anti-tetanus vaccine through rational design of genetically modified protein and also in designing structure-based small molecule inhibitors.

Acknowledgements

Research supported by the US Army Medical Research Acquisition Activity (Award No. DAMD17-02-2-0011) under DOE Prime Contract No. DE-AC02-98CH10886 with Brookhaven National Laboratory. T.B. was supported by a grant RGY0027/2001 from Human Frontier Science Program. We would like to thank Drs A. Saxena and M. Becker for providing beamtime at the NSLS and Drs R. Agarwal, S. Eswaramoorthy and S. Jayaraman for helpful discussions.

References

Agarwal, R., Eswaramoorthy, S., Kumaran, D., Binz, T., Swaminathan, S., 2004a. Structural analysis of botulinum neurotoxin type E catalytic domain and its mutant Glu212→Gln reveals the pivotal role of the Glu212 carboxylate in the catalytic pathway. Biochemistry 43, 6637–6644.

Agarwal, R., Eswaramoorthy, S., Kumaran, D., Dunn, J.J., Swaminathan, S., 2004b. Cloning, high level expression,

purification and crystallization of the full length *Clostridium botulinum* neurotoxin type E light chain. Protein Expr. Purif. 34, 95–102.

Binz, T., Blasi, J., Yamasaki, S., Baumeister, A., Link, E., Sudhof, T.C., Jahn, R., Niemann, H., 1994. Proteolysis of SNAP-25 by Types E and A botulinum neurotoxins. J. Biol. Chem. 269, 1617–1620.

Binz, T., Bade, S., Rummel, A., Kollewe, A., Alves, J., 2002. Arg^{362} and Tyr^{365} of the botulinum neurotoxin type A light chain are involved in transition state stabilization. Biochemistry 41, 1717–1723.

Breidenbach, M.A., Brunger, A., 2004. Substrate recognition strategy for botulinum neurotoxin serotype A. Nature 432, 925–929.

Brunger, A.T., Adams, P.D., Clore, G.M., Delano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszwewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Somonsom, T., Warren, G.L., 1998. Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D54, 905–921.

Burley, S.K., Petsko, G.A., 1986. Amino–aromatic interactions in proteins. FEBS Lett. 203, 139–143.

Cornille, F., Martin, L., Lenoir, C., Cussac, D., Roques, B.P., Fourine-Zaluski, M., 1997. Cooperative exosite dependent cleavage of synapotbrevin by tetanus toxin light chain. J. Biol. Chem. 272, 3459–3464.

Cowtan, K., 1994. Joint CCP4 ESF-EACBM Newslett. Protein Crystallogr. 31, 34–38.

De La Fortelle, E., Bricogne, G., 1997. Maximum-likelihood heavy atom parameter refinement in the MIR and MAD methods. Meth. Enzymol. 276, 472–493.

Emsley, P., Fotinou, C., Black, I., Fairweather, N.F., Charles, I.G., Watts, C., Hewitt, E., Isaacs, N.W., 2000. The structures of the H(C) fragment of tetanus toxin with carbohydrate subunit complexes provide insight into ganglioside binding. J. Biol. Chem. 275, 8889–8894.

Eswaramoorthy, S., Kumaran, D., Keller, J., Swaminathan, S., 2004. Role of metals in the biological activity of *Clostridium botulinum* neurotoxins. Biochemistry 43, 2209–2216.

Fillippis, V.D., Vangelista, L., Schiavo, G., Tonello, F., Montecucco, C., 1995. Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin. Eur. J. Biochem. 229, 61–69.

Foran, P., Shone, C.C., Dolly, J.O., 1994. Differences in the protease activities of tetanus and botulinum B toxins revealed by the cleavage of vesicle-associated membrane protein and various sized fragments. Biochemistry 33, 15365–15374.

Galazka, A., Gasse, F., 1995. The present status of tetanus and tetanus vaccination. Curr. Top. Microbiol. Immunol. 195, 31–53.

Hanson, M.A., Stevens, R.C., 2000. Cocrystal structure of synaptobrevin-ll bound to botulinum neurotoxin type B at 2.0 A resolution. Nat. Struct. Biol. 7, 687–692.

Jankovic, J., Hallett, M., 1994. Therapy with Botulinum Toxin. Marcel Dekker, New York.

Jones, T.A., Zou, J., Cowtan, S., Kjeldgaard, M., 1991. Improved methods in building protein models in electron density map and the location of errors in these models. Acta Crystallogr. A47, 110–119.

Kleywegt, G.J., Jones, T.A., 1997. Detecting folding motifs and similarities in protein structures. Meth. Enzymol. 277, 525–545.

Kurazono, H., Mochida, S., Binz, T., Eisel, U., Quanz, M., Grebenstein, O., Wernars, K., Poulain, B., Tauc, L., Niemann, H., 1992. Minimal essential domains specifying toxicity of the light chains of tetanus toxin and botulinum neurotoxin type A. J. Biol. Chem. 267, 14721–14729.

Lacy, D.B., Stevens, R.C., 1998. Unravelling the structures and modes of action of bacterial toxins. Curr. Opin. Struct. Biol. 8, 778–784.

Lacy, D.B., Stevens, R.C., 1999. Sequence homology and structural analysis of clostridial neurotoxins. J. Mol. Biol. 291, 1091–1104.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality for assessing the accuracy of protein structures. J. Appl. Crystallogr. 26, 283–291.

Li, Y., Foran, P., Fairweather, N.F., de Paiva, A., Weller, U., Dougan, G., Dolly, J.O., 1994. A single mutation in the recombinant light chain of tetanus toxin abolishes its proteolytic activity and removes the toxicity seen after reconstitution with native heavy chain. Biochemistry 33, 7014–7020.

McMahon, H.T., Ushkaryov, Y.A., Edelmann, L., Link, E., Binz, T., Niemann, H., Jahn, R., Sudhof, T.C., 1993. Cellubrevin is a ubiquitous tetanus-toxin substrate homologous to a putative synaptic vesicle fusion protein. Nature 364, 346–349.

Montecucco, C., Schiavo, G., 1995. Structure and function of tetanus and botulinum neurotoxins. Q. Rev. Biophys. 28, 423–472.

Navaza, J., Saludjian, P., 1997. AMoRe: an automated molecular replacement program package. Meth. Enzymol. 276, 581–594.

Otwinowski, Z., Minor, W., 1997. Processing of X-ray diffraction data collected in oscillation mode. Meth. Enzymol. 276, 307–326.

Pellizzari, R., Rossetto, O., Lozzi, L., Giovedi', S., Johnson, E., Shone, C.C., Montecucco, C., 1996. Structural determinants of the specificity for synaptic vesicle-associated membrane protein/synaptobrevin of tetanus and botulinum type B and G neurotoxins. J. Biol. Chem. 271, 20353–20358.

Rappuoli, R., Podda, A., Pizza, M., Covacci, A., Bartoloni, A., de Magistris, M.T., Nencioni, L., 1992. Progress towards the development of new vaccines against whooping cough. Vaccine 10, 1027–1032.

Rawlings, N.D., Barrett, A.J., 1995. Evolutionaly families of metallopeptidases. Meth. Enzymol. 248, 183–228.

Rigoni, M., Caccin, P., Johnson, E.A., Montecucco, C., Rossetto, O., 2001. Site-directed mutagenesis identifies active-site residues of the light chain of botulinum neurotoxin type A. Biochem. Biophys. Res. Commun. 288, 1231–1237.

Rossetto, O., Caccin, P., Rigoni, M., Tonello, F., Bortoletto, N., Stevens, R.C., Montecucco, C., 2001. Active-site mutagenesis of tetanus neurotoxin implicates TYR-375 and GLU-271 in metalloproteolytic activity. Toxicon 39, 1151–1159.

Schiavo, G., Santucci, A., Dasgupta, B.R., Metha, P.P., Jontes, J., Benfenati, F., Wilson, M.C., Montecucco, C., 1993. Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct COOH-terminal peptide bonds. FEBS Lett. 335, 99–103.

Schiavo, G., Rossetto, O., Benfenati, F., Poulain, B., Montecucco, C., 1994. Tetanus and botulinum neurotoxins are zinc proteases specific for components of the neuroexocytosis apparatus. Ann. NY. Acad. Sci. 710, 65–75.

Schiavo, G., Matteoli, M., Montecucco, C., 2000. Neurotoxins affecting neuroexocytosis. Physiol. Rev. 80, 717–766.

Segelke, B., Knapp, M., Kadkhodayan, S., Balhorn, R., Rupp, B., 2004. Crystal structure of *Clostridium botulinum* neurotoxin protease in a product-bound state: evidence for noncanonical zinc protease activity. Proc. Natl Acad. Sci. USA 101, 6888–6893.

Simpson, L.L., 1986. Molecular pharmacology of botulinum toxin and tetanus toxin. Annu. Rev. Pharmacol. Toxicol. 26, 427–453.

Simpson, L.L., Maksymowych, A.B., Hao, S., 2001. The role of zinc binding in the biological activity of botulinum toxin. J. Biol. Chem. 276, 27034–27041.

Swaminathan, S., Eswaramoorthy, S., 2000. Structural analysis of the catalytic and binding sites of *Clostridium botulinum* neurotoxin B. Nat. Struct. Biol. 7, 693–699.

Swaminathan, S., Eswaramoorthy, S., Kumaran, D., 2004. Structure and enzymatic activity of botulinum neurotoxins. Mov. Disord. 19 (Suppl. 8), S17–S22.

Terwilliger, T.C., Berendzen, J., 1997. Automated structure solution for MIR and MAD. Acta Crystallogr. D55, 849–861.

Tonello, F., Pellizzari, R., Pasqualato, S., Grandi, G., Peggion, E., Montecucco, C., 1994. Recombinant and truncated tetanus neurotoxin light chain: cloning, expression, purification, and proteolytic activity. Protein Expr. Purif. 15, 221–227.

Umland, T.C., Wingert, L.M., Swaminathan, S., Furey, W.F., Schmidt, J.J., Sax, M., 1997. Structure of the receptor binding fragment H_c of tetanus neurotoxin. Nat. Struct. Biol. 4, 788–792.

Vaidyanathan, V.V., Yoshino, K., Jahnz, M., Dorries, C., Bade, S., Nauenburg, S., Niemann, H., Binz, T., 1999. Proteolysis of SNAP-25 isoforms by botulinum neurotoxin types A, C, and E: domains and amino acid residues controlling the formation of enzyme–substrate complexes and cleavage. J. Neurochem. 72, 327–337.

Washbourne, P., Pellizzari, R., Baldini, G., Wilson, M.C., Montecucco, C., 1997. Botulinum neurotoxin A and E require the SNARE motif in SNAP-25 for proteolysis. FEBS Lett. 418, 1–5.

Yamasaki, S., Hu, Y., Binz, T., Kalkuhl, A., Kurazono, H., Tamura, T., Jahn, R., Kandel, E., Niemann, H., 1994. Synaptobrevin/vesicle-associated membrane protein (VAMP) of *Aplysia californica*: structure and proteolysis by tetanus toxin and botulinal neurotoxins type D and F. Proc. Natl Acad. Sci. USA 91, 4688–4692.